Self-Organization of Hierarchical Visual Maps with Feedback Connections

Yiu Fai Sit and Risto Miikkulainen

Department of Computer Sciences
The University of Texas at Austin
Austin, TX 78712, U.S.A.

Abstract

Visual areas in primates are known to have reciprocal connections. While the feedforward bottom-up processing of visual information has been studied extensively for decades, little is known about the role of the feedback connections. Existing feedback models usually employ hand-coded connections, and do not address how these connections develop. The model described in this paper shows how feedforward and feedback connections between cortical areas V1 and V2 can be learned through self-organization simultaneously. Computational experiments show that both areas can form hierarchical representations of the input with reciprocal connections that link relevant cells in the two areas.

Key words: feedback connections, self-organization, visual cortex

---

1 Introduction

Reciprocal connections are very common between two different areas in the visual cortex of primates [6, 10]. Most of the previous studies of perception in primates, however, only concentrate on feedforward connections that link neurons from a lower level, such as V1, to a higher level, such as V2. While a lot of progress has been made in understanding the function of the feedforward connections, the development and the role of feedback connections in the visual system remain unclear.

Recently, several models of primate visual system with feedback connections have been proposed [2–4, 9]. In these models, either the receptive fields or the weights of feedback connections are hand-coded in some way. Although the behavior of these models matches biological data, it is unclear whether the hand-coded connection patterns are biologically plausible and how the visual system comes up with them during development.

This paper shows how reciprocal connections between V1 and V2 can be developed through self-organization. With corners composed of two line segments as input, V1 and V2 form hierarchical representations of these patterns via self-organization.

Email address: {`yfsit, risto`}`@cs.utexas.edu` (Yiu Fai Sit and Risto Miikkulainen).

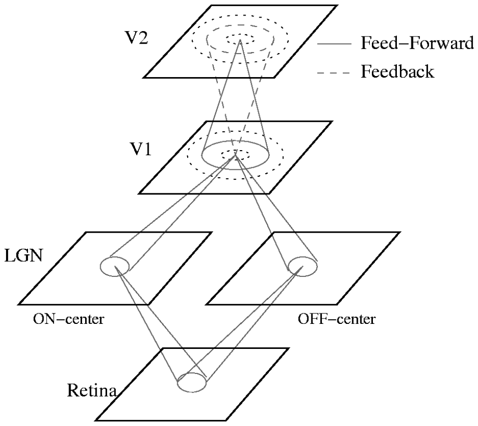

Fig. 1. Extended LISSOM model with V2 and feedback connections. The model is a hierarchy of sheets of neural units that represent the retina, ON-center and OFF-center LGN, V1, and V2. Sample connections are indicated for a unit in each sheet. The feedforward connections to an LGN neuron form a receptive field on the retina, and V1 and V2 neurons have receptive fields on the LGN and V1 sheets, respectively. In the experiments reported in this paper, however, the LGN was bypassed for simplicity; V1 receives input directly from the retina, making the model faster to simulate with equivalent results [8]. V1 and V2 neurons also have short-range lateral excitatory (small dotted circle) and long-range lateral inhibitory (large dotted circle) connections. In addition to these connections, V1 neurons receive feedback connections from V2 neurons that form a receptive field on V2. V2 neurons have a larger receptive field than V1 neurons, as in the visual system of primates.

Feedback connections link units in V2 to V1 that are activated by the same input, demonstrating for the first time that meaningful feedback connections can be formed between two areas by self-organization.

2 Extended LISSOM Model

The experiments are based on an extension of the LISSOM (Laterally Interconnected Synergetically Self-Organizing Map) model [1, 8, 13] to include feedback connections from V2 to V1. The feedback connections are excitatory as suggested by some studies [11, 12]. Figure 1 shows the architecture of this extended model.

The software implementation of the model is largely based on [1]. Since the focus of this study is on the organization between V1 and V2, the processing at LGN is bypassed to speed up the computation. With artificial inputs like elongated Gaussians, LGN can safely be skipped to produce qualitatively equivalent result in the LISSOM model [8]. Each neural unit in the map corresponds to a column of cells through the six anatomical layers of the cortex.

For each input, the initial response $\eta_a(0)$ for the V1 unit a is

$$\eta_a(0) = \sigma_1(\sum_r \chi_r A_{r,a}),$$

where σ_1 is a piecewise linear sigmoid activation function for the V1 map, χ_r is the activation of the retinal unit r, and $A_{r,a}$ is the afferent weight value that connects the unit r to the unit a. The initial responses of V1 are input to the V2 units b, which in turn produce initial response $\eta_b(0)$:

$$\eta_b(0) = \sigma_2(\sum_a \eta_a(0) A_{a,b}).$$

After the initial activation of V2, V1 activity settles through short-range excitatory and long-range inhibitory lateral interactions, as well as excitatory feedback from V2:

$$\eta_a(t) = \sigma_1(\sum_r \chi_r A_{r,a} + \sum_\rho \gamma_\rho \sum_{\hat{a}} \eta_{\hat{a}}(t-1) L_{\rho\hat{a},a} + \sum_b \eta_b(t-1) F_{b,a}),$$

where the first term is the weighted sum of the activations from the retina, the second term is the sum of lateral activations in V1 during the previous time step, and the last term is the weighted sum of the activations in V2 during the previous time step. The label ρ in the second term identifies the type (excitatory or inhibitory) of lateral connection weights L, and γ_ρ is a constant scaling factor for each ρ (negative for inhibitory connections). The index $\hat{a}$ spans the whole V1 map. $F_{b,a}$ is the feedback connection weight from V2 unit b to V1 unit a. The activation of V2 is given by a similar equation, but without the feedback term:

$$\eta_b(t) = \sigma_2(\sum_a \eta_a(t) A_{a,b} + \sum_\rho \gamma_\rho \sum_{\hat{b}} \eta_{\hat{b}}(t-1) L_{\rho\hat{b},b}).$$

These equations assume that feedforward and feedback connections are twice as fast as lateral connections, which is an approximation of biological visual systems [7].

After the activities have settled, the connection weights of all V1 and V2 units are modified according to the Hebb rule, with normalization. The lateral connection weights to a V1 unit a are updated as

$$L'_{\rho\hat{a},a} = \frac{L_{\rho\hat{a},a} + \alpha_\rho \eta_{\hat{a}} \eta_a}{\sum_{\hat{a}} (L_{\rho\hat{a},a} + \alpha_\rho \eta_{\hat{a}} \eta_a)},$$

where $L'_{\rho\hat{a},a}$ is the new connection weight, α_ρ is the learning rate for each type of lateral connections. The lateral connection weights in V2 are modified analogously.

The feedforward and feedback connection weights of a V1 unit are normalized together as one group. While it is possible to treat them separately, the appropriate proportions of feedforward and feedback contributions to V1 activation would have to be set by hand. By normalizing these two connection types together, these proportions self-organize:

$$\begin{aligned} A'_{r,a} &= \frac{A_{r,a} + \alpha \chi_r \eta_a}{\sum_a (A_{r,a} + \alpha \chi_r \eta_a) + \sum_b (F_{b,a} + \alpha \eta_b \eta_a)} \\ F'_{b,a} &= \frac{F_{b,a} + \alpha \eta_b \eta_a}{\sum_a (A_{r,a} + \alpha \chi_r \eta_a) + \sum_b (F_{b,a} + \alpha \eta_b \eta_a)}. \end{aligned}$$

For a V2 unit b, the afferent connection weights from V1 are updated as

$$A'_{a,b} = \frac{A_{a,b} + \alpha \eta_a \eta_b}{\sum_a (A_{a,b} + \alpha \eta_a \eta_b)}.$$

3 Experiments

Some V2 cells in the visual cortex of primates are known to be selective to corner-like patterns [5]. It is therefore reasonable to use such simple patterns as a first step to study the behavior of the extended LISSOM model. In the experiment, right-angled corners composed of two edges were used as the input. Four corner orientations were included: $<, \wedge, >$, and $\vee$. At each learning step, one orientation was selected in a fixed order and the corner was placed at a random position in the retina.

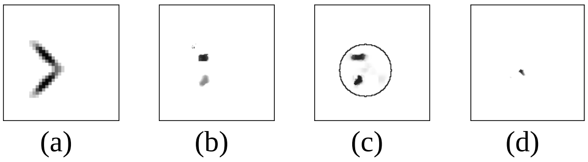
(a) (b) (c) (d)

Fig. 2. A sample input and responses in V1 and V2. (a) An input corner composed of two elongated Gaussians shown as activation values from black (high) to white (low) on the retina. (b) Response of the organized V1 map, intensity-coded to represent the orientation preferences of the V1 units that are activated, with light gray representing 45^o and dark gray -45^o. The V1 map forms a representation of the input corner as two separate lines. (c) Receptive field of a V2 unit that is activated by the input pattern with weight values shown in gray scale from black to white. The receptive field matches the activity in V1 and thus is selective to corner input of this orientation at this position. The black circle shows the boundary of the receptive field. (d) Output of the V2 map in gray scale from black to white. The activation represents a corner at this location; there are no separate responses to the edges. A hierarchical representation of the corner is therefore formed in V1 and V2.

After the responses in V1 and V2 had settled, both of the maps updated the connection weights of all units. Figure 2 shows an example input and the corresponding responses in V1 and V2 after self-organization.

The retina had 36×36 units, while both V1 and V2 maps had 144×144 units. The receptive field radii of V1 and V2 units were 5 and 24, respectively, modeling the human visual system. All weights were initialized randomly, except that the total feedback weights to V1 were half of the total feedforward weights from the retina initially.

4 Experimental Results

After $20,000$ input presentations, V1 map self-organized and stabilized to form units that were highly selective to edges (Figure 3a). The resulting V1 map (Figure 4a) is very similar to those obtained in similar studies where edges were used as inputs [8].

Most of the V2 units formed receptive fields that were selective to the patterns of V1 activations that corresponded to a corner. As in the V1 map, units that preferred similar patterns clustered together (Figure 4b).

Feedback connections self-organized to link those V2 and V1 units that were both active for a particular input corner and hence reciprocal connections had developed. Since an edge in one position can be a part of several different corners, V1 unit received feedback from many V2 units at different positions (Figure 3b).

5 Discussion

In this model, hierarchical representation of corners in V1 and V2 emerges because these areas have different receptive field sizes. A receptive field in V1 is only large enough to cover part of an edge of a corner. A receptive field can also cover the tip of the corner that contains parts of both edges, but this happens much more rarely. Over a large number of input presentations, V1 receptive fields encounter more edges than the tips and develop edge selectivity through Hebbian learning (Figure

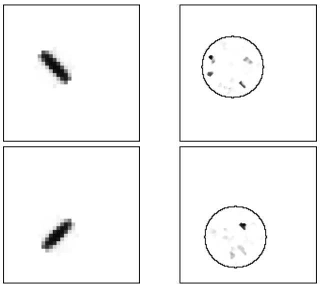
(a) (b)

Fig. 3. The receptive fields and feedback connections of two active V1 units in Figure 2 represented in gray scale coding. (a) V1 units form edge-selective receptive fields even though the inputs are corners. (b) Feedback connections received from V2 units for each of the V1 units in (a). Since an edge in the input can be a part of several different corners, V1 units receive feedback from patches of V2 that represent different corners that share the same edge. For example, the four main clusters in the top figure consist of four groups of V2 units, each representing a corner that shares the edge. Both V1 units receive feedback from the V2 unit in Figure 2d (the bottom right patch in the top figure and the top right patch in the bottom figure). These results show that feedback connections organize to link units in V1 and V2 that are part of the representation for the same input pattern.

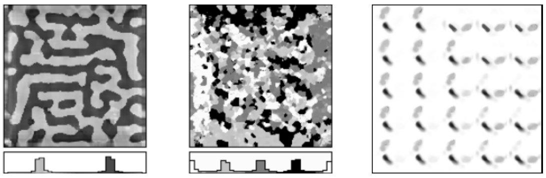
(a) (b) (c)

Fig. 4. Orientation preferences in the V1 map, corner preferences in the V2 map, and a sample of V2 receptive fields. (a) The intensity in the V1 map represents the preference of the unit at that point: light gray is 45^o and dark gray is -45^o. The histogram at the bottom shows the distribution of orientation preference in the map. Similar orientation maps are formed by using elongated Gaussian inputs with no feedback. (b) The orientation of the corner that a V2 unit prefers is intensity-coded: white represents $>$, light gray $\wedge$, dark gray $<$, and black $\vee$. The histogram shows that the four orientations are represented evenly in the map. The large patches of same preference on the borders are due to edge effects. (c) The receptive fields of 25 V2 units around the center region of the map, using the same intensity code in (a). Units that prefer similar patterns form clusters. The gradual change in receptive fields across the clusters is very similar to that in the V1 map, showing that the hierarchical organization follows the same principles at both levels.

4a). On the other hand, the receptive fields in V2 are large enough to cover the activity patterns in V1 for both edges and hence V2 units become corner selective. Although the inputs are artificial and simple in the current experiment, similar results should be obtained when more complex inputs like natural images are used [8].

Consistent with biology, feedback is relatively weak compared to the input signal [11]. As a result, it does not significantly influence how V1 develops in this model; whether it does in biology is currently an open question. However, feedback is likely to play a large role in visual function, including perceptual grouping such as contour completion, and many visual illusions. The current model is well-suited for studying such effects in the future.

6 Conclusions

The computational results in this paper showed for the first time that afferent, lateral, and feedback connections may all self-organize simultaneously using the same normalized Hebbian learning principle to form hierarchical representation of the inputs. An immediate future step is to use natural images as input and match the result with biological data on V2 and its feedback connections. The model can also be used to confirm existing hypotheses and form new ones about the role of feedback connections in vision, in particular in perceptual grouping. Another possible direction is to model multi-modal interactions of different functional areas. Understanding self-organization of a feedforward and feedback hierarchy is therefore an important step towards a more general understanding of the dynamics of the visual system.

Acknowledgments

This research was supported in part by the NIMH under Human Brain Project grant 1R01-MH66991 and NSF under grant EIA-0303609.

References

[1] J. A. Bednar, *Learning to See: Genetic and Environmental Influences on Visual Development*, Ph.D. thesis, Department of Computer Sciences, The University of Texas at Austin, Austin, TX (2002), Technical Report AI-TR-02-294.

[2] G. Deco and T. S. Lee, The role of early visual cortex in visual integration: A neural model of recurrent interaction, *European Journal of Neuroscience* (2004), 20:1089–1100.

[3] G. Deco and E. T. Rolls, A neurodynamical cortical model of visual attention and invariant object recognition, *Vision Research* (2003), 44:621–642.

[4] T. Hansen and H. Neumann, Neural mechanisms for the robust representation of junctions, *Neural Computation* (2004), 16:1013–1037.

[5] J. Hegdé and D. V. Essen, Selectivity for complex shapes in primate visual area V2, *Journal of Neuroscience* (2000), 20:RC61:1–6.

[6] L. C. Katz and E. M. Callaway, Development of local circuits in mammalian visual cortex, *Annual Review of Neuroscience* (1992), 15:31–56.

[7] V. A. Lamme, H. Supèr, and H. Spekreijse, Feedforward, horizontal, and feedback processing in the visual cortex, *Current Opinion in Neurobiology* (1998), 8:529–535.

[8] R. Miikkulainen, J. A. Bednar, Y. Choe, and J. Sirosh, *Computational Maps in the Visual Cortex* (Springer-Verlag, Berlin; New York, 2005), in press.

[9] H. Neumann and W. Sepp, Recurrent V1-V2 interaction in early visual boundary processing, *Biological Cybernetics* (1999), 81:425–444.

[10] P. Salin and J. Bullier, Corticocortical connections in the visual system: structure and function, *Physiological Review* (1995), 75:107–154.

[11] J. Sandell and P. Schiller, Effect of cooling area 18 on striate cortex cells in the squirrel monkey, *Journal of Neurophysiology* (1982), 48:38–48.

[12] Z. Shao and A. Burkhalter, Different balance of excitation and inhibition in forward and feedback circuits of rat visual cortex, *Journal of Neuroscience* (1996), 16:7353–7365.

[13] J. Sirosh, R. Miikkulainen, and J. A. Bednar, Self-organization of orientation maps, lateral connections, and dynamic receptive fields in the primary visual cortex, in: J. Sirosh, R. Miikkulainen, and Y. Choe, eds., *Lateral Interactions in the Cortex: Structure and Function* (The UTCS Neural Networks Research Group, Austin, TX, 1996), electronic book, ISBN 0-9647060-0-8, http://nn.cs.utexas.edu/web-pubs/htmlbook96.