Recurrent on-center off-surround neural networks and human visual system: a computational framework

Rakesh Sengupta
Center for Cognitive Sciences, BVRIT-Hyderbad, India

Anindya Pattanayak
Viswavarati University, India

Dipanjan Roy
Center for Cognitive Sciences, IIIT-Hyderabad, India

S. Bapiraju
Center for Cognitive Sciences, IIIT-Hyderabad, India

Abstract

Human cortical structure is comprised of dense feedforward connection between layers (through pyramidal cells) and recurrent pool of excitatory and inhibitory connections within layers. Human visual system for instance has feedforward activity in early visual areas from V1 to V4 (with feedback loop from higher areas) that goes into parietal and temporal cortex where the signals interact with the recurrent activity in those layers as well as working memory networks in dl-PFC and even long term memory networks (Itti & Koch, 2001). Computationally the recurrent cortical connectivity can be represented through two variants of the Cohen-Grossberg networks (Grossberg, 2013, 1973): additive and shunting on-center off-surround architecture. We have two goals in the current paper. Firstly, we intend to show how these two kinds of networks can be used to explain several different cognitive phenomena beyond the ones available in state-of-the-art literature. Secondly, we introduce a novel energy function that can take us from brain to behavior. We show application of the network in case of visual numbers, working memory, and individuation.

Network architechture

The additive variant of the network was introduced in our previous work Sengupta, Bapi Raju, and Melcher (2014). The network uses a single layer of N completely connected self-excitatory nodes that laterally inhibit each other. The strengths of self-excitation and lateral inhibition are given by α and β respectively. The temporal dynamics of the system is given by the following equation,

$$\dot{x}_i = \frac{dx_i}{dt} = -x_i + \alpha F(x_i) - \beta \sum_{j=1, j\neq i}^{N} F(x_j) + I_i + noise \quad (1)$$

where x_i is the activation of i-th node. I_i represents the transient input to the i-th node that lasts for finite time and has constant value in the interval $[0,1]$ in the duration. $F(x)$ is the activation function given by

$$F(x) = \begin{cases} 0 & \text{for } x \leq 0 \\ \frac{x}{1+x} & \text{for } x > 0 \end{cases} \quad (2)$$

The shunting network on the other hand is described by additional multiplicative terms and two parameters that form Shunting range $[-D, B]$. The network dynamics varies with the equation:

$$\frac{dx_i}{dt} = -Ax_i + (B - x_i)F(x_i) - (D + x_i) \sum_{k=1, k\neq i}^{N} F(x_k) + I_i + noise \quad (3)$$

In our previous work (Sengupta et al., 2014) we took a novel approach towards characterizing network stability in terms of a Hamiltonian function. It had the advantage of explain the analytic solution space of the network in more natural terms, as well as derive constraint equations that were not visible from network analysis alone, but could be confirmed through simulations (see Appendix of Sengupta et al. (2014) for detailed derivation and further explanations). The additive network has energy given by

$$H_{\text{additive}} = \sum_i H_i \propto -\sum_i \int \left(1 - \alpha \left(\frac{F(x_i)}{x_i}\right)^2\right) \dot{x}_i^{\,2} dt \quad (4)$$

$\dot{x}_i = \frac{dx_i}{dt}$ and H_i is the energy for a particular node i.

$$H_{\text{shunting}} = \sum_i H_i \propto -\sum_i \int \left(1 + \sum_{i=1}^{N} F(x_i) - (B - x_i)\left(\frac{F(x_i)}{x_i}\right)^2\right) \dot{x}_i^{\,2} dt \quad (5)$$

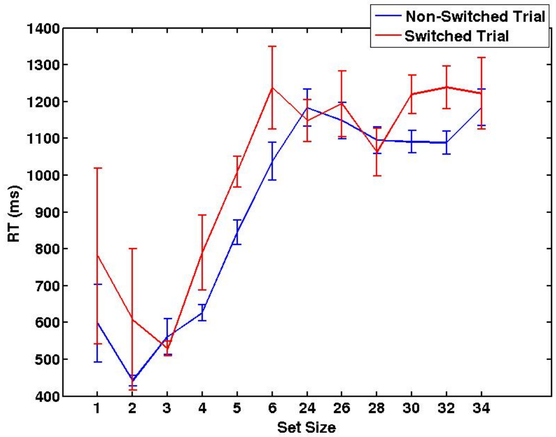
We could formulate the reaction times for certain tasks could be reformulated as the following.

$$RT \sim -\sum H_i \quad (6)$$

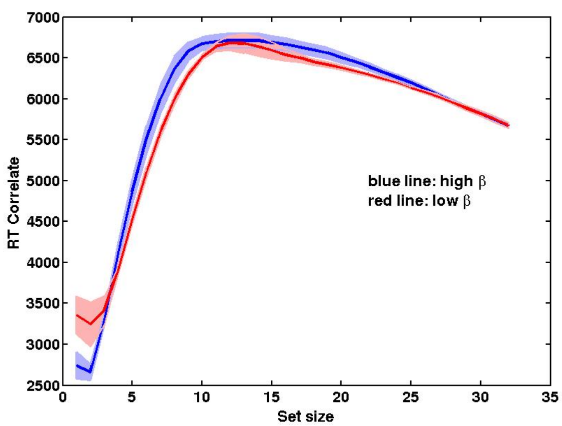
The idea behind such formulation was that the behavioral reaction time would be highly correlated with the maximum allowed fluctuation of energy so that the network goes back to the original state.

Results

We have used the model to explain a variety of phenomena. For instance the additive model explains the reaction

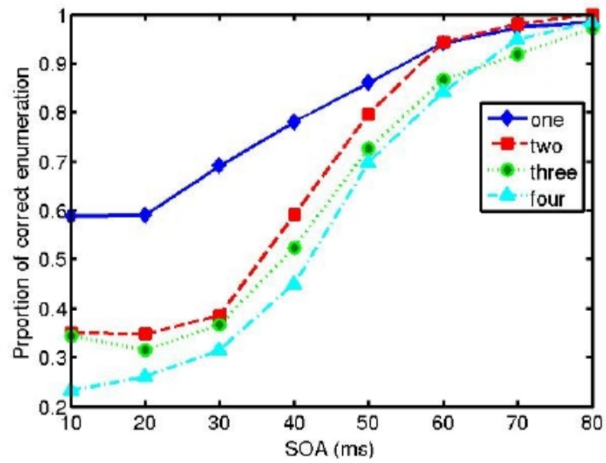


(a) Average reaction times for each numerosity plotted with error bars for standard/non-switched and switched trials. Non-switch trials refer to when a small numerosity trial followed a small numerosity trial or large numerosity trial followed a large numerosity trial

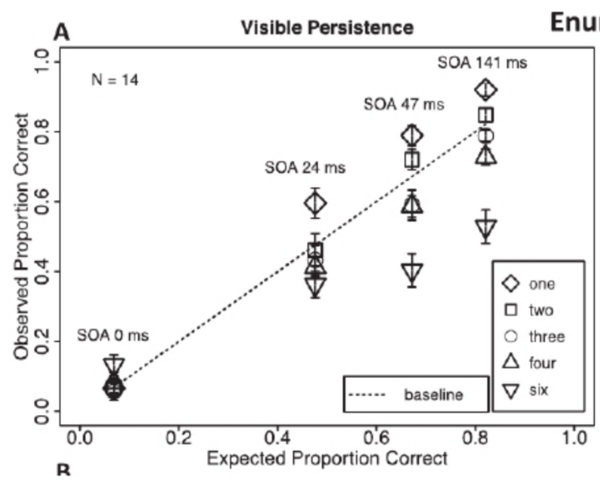


(b) This figure shows the average RT correlate or RT_{model} values for two ranges of neural inhibition parameter β: high and low. The higher numerosity ranges show similar patterns of RT for the two ranges of inhibition, while the low numerosity range differs greatly in predicted RT between the two beta ranges. The error bands represent one standard deviation of the mean RT_{model} values at different set sizes.

Figure 1. Empirical and model results for enumeration of different number of items



(a) Model results for individuation performance in a forward mask design with different SOAs



(b) Empirical results for individuation with different SOAs in forward masking paradigm (reproduced with permission from Wutz and Melcher, 2014

Figure 2. Additive network performance for individuation at different set sizes

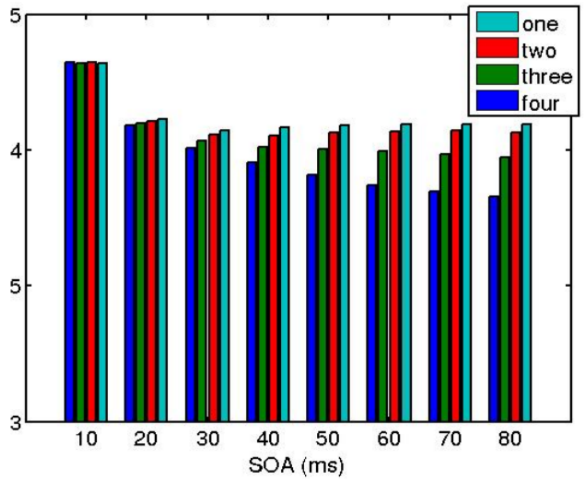
times for enumeration at different time scales as shown in the figure below (Fig. 1). In the experiment a series of 200 trials were presented. The trials consisted of display of black dots persisting for 100 ms and participants verbally reported the number of items as quickly and accurately as possible. The white flash was used as a mask to reduce the visual persistence of the dots in order to encourage participants to respond quickly rather than slowly counting each item individually from the after-image.

The additive variant also explains differences in individuation performance for different set sizes and their corresponding reaction times (Fig. 2 and 3.

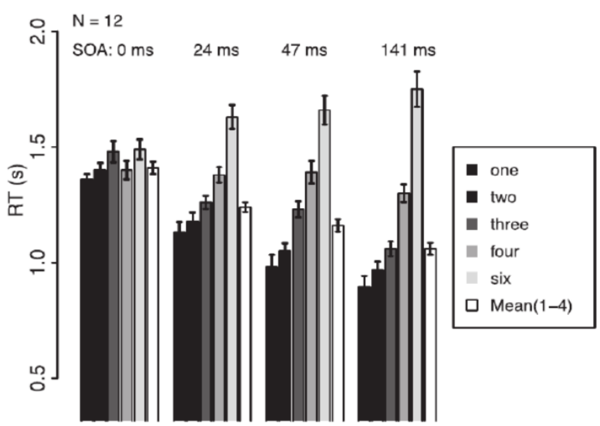
The shunting variant of the network shows interesting results for texture processing of visual objects in a crowded display.

Conclusions

The current work suggests that the recurrent on-center off-surround activity can give a clue towards various phenomena in visual domain including visual working memory, individuation and visual sense of numbers. In fact a surround-suppression mechanism is ubiquitous throughout the visual



(a) RT for individuation from model results calculated using the RT-correlate of energy function in 6.



(b) RT in empirical experiment for different SOAs. Reproduced with permission from Wutz and Melcher, 2014

Figure 3. Reaction times for the individuation paradigm from model and empirical results

system. In the current work we have shown that we can connect the network properties to directly observable behavioral results like reaction times with the help of a novel energy function. We plan to investigate the properties of such networks in more detail in future.

References

Grossberg, S. (1973). Contour enhancement, short term memory, and constancies in reverberating neural networks. *Studies in Applied Mathematics*, *LII*, 213-257.

Grossberg, S. (2013). Recurrent neural networks. *Scholarpedia*, *8*(2), 1888.

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, *2*, 194-203.

Sengupta, R., Bapi Raju, S., & Melcher, D. (2014). A visual sense of number emerges from the dynamics of a recurrent on-center off-surround neural network. *Brain Research*, *1582*, 114-124.