

Image Segmentation Using Neural Oscillators

Talia Konkle¹, Ning Jiang², Jie Zhang³, Fatma Gurel⁴,
Christopher Scheper⁵ and Gheorghe Craciun⁶

¹ Institute for Cognitive and Brain Sciences and Department of Mathematics, University of California at Berkeley, ² Institute of Biomedical Engineering, University of New Brunswick, ³ Department of Applied Science, The College of William and Mary, ⁴ Department of Mathematics, University of Pittsburgh, ⁵ Department of Mathematics and Department of Physics, Purdue University, ⁶ Mathematical Biosciences Institute, Ohio State University

Abstract

Synchrony in neural oscillations is an observed phenomenon in neurobiology and has been proposed as a possible neural mechanism addressing the binding problem (Malsburg and Buhmann, 1992). We attempted to mathematically model this binding synchrony as it pertained to a two-dimensional binary visual input -- more simply, neurons corresponding to an object A should fire together, and these neurons should fire out of phase with neurons corresponding to object B . We used reduced Hodgkin-Huxley equations for our neuronal model, eliminating the spatial dimension in order to work with ordinary differential equations. We added 'synchrony terms' to these equations accounting for the synaptic connectivity, changing the strength and type of these connections in order to achieve object binding and image segmentation. Using nearest-neighbor excitatory connections in the sensory (input) neuronal layer, and introducing a global inhibitor receiving from and transmitting to every sensory cell (Terman and Wang, 1995), we found that this connectivity would produce the desired results when a proper balance of the synaptic conductance parameters was found. We briefly examined further extensions to this model by incorporating moving objects.

This material is based upon work supported by the National Science Foundation under Agreement No. 0112050.

Introduction

The neurobiology of the visual system is the most studied sensory system to date, so there is much known about the circuitry and levels of processing of a visual signal (Kandel et al., 2000). The first stage of processing occurs in the retina of the eyes. Here photons excite the receptor cells (rods and cones), which converge to the bipolar cells and ganglion cells. The ganglion cell axons exit the eye and make up the optic nerve. Since information in the brain is processed contralaterally, the input from the right and left visual fields and the right and left eyes must sort themselves out so that all the information from the right side of space gets sent to the left hemisphere, and visa versa. This ‘sorting out’ occurs when the optic nerves cross at the optic chiasm.

These axons then synapse onto cells in the lateral geniculate nucleus (LGN). Here there is a little more convergence in the visual signal; however, the connections are roughly one-to-one, so the form of the visual signal is not modified dramatically. About 10% of these cells send axons to the superior colliculus and the supra chiasmatic nucleus, which are involved in eye movements and the regulation of biological clocks, respectively. The other 90% project back to the primary visual cortex (V1) in the occipital lobe. Here there is further convergence and processing of the visual signal.

To understand the processing of this visual signal, we compare the responses of cells in retina and in V1 given a simple visual stimulus. Cell recordings in the retina show that cells fire maximally when there is a single point of light shining in a cell’s small receptive field. However, when presenting this same point of light while recording from cells in V1, no cell will fire very strongly. Instead, cells in V1 fire maximally when a bar of light is presented at a certain orientation. This demonstrates that the convergence of cells corresponds to a more processed visual signal. From V1, the visual signal is sent to other “visual maps” important for different features of the visual image (e.g. motion and color) for further refined processing.

The binding problem

This leads us directly to the so-called ‘binding problem’, which can be stated in a simplified manner as follows: if different parts of the brain are processing different features of the *same* object, how are these features bound together so that we have a coherent unified mental construction of that object? Concretely, if we see a fast-moving blue object and a red octagonal shiny object, we are not often mistaken that it’s really a red moving object and a blue shiny object. One obvious question that arises then, is what neural mechanism allows us to bind together the proper features of an object, and further, to keep those features separate from the features of another object? (Note that these different components are not limited to the visual features but also to auditory features and tactile features etc. However, for simplicity, we will only examine the binding problem as it pertains to the visual domain.)

In the neurobiological literature, there is evidence of synchronous oscillations of neuron firing. In the cat visual cortex, for example, single cell recordings have shown synchronous oscillatory behavior, even across distant areas (Eckhorn et al., 1988.) When recording from different areas in V1, if distant cells have similar receptive fields, meaning that they fire maximally to the same stimulus, then it is observed that their firing is in sync, and it is at a roughly steady frequency.

These scientific observations have led to the proposal of neural synchrony as a possible solution to the binding problem. Neurons which are processing a signal from one object, regardless of where they are in the brain, will fire in sync with each other. By extension then, they will fire out of sync with neurons processing information about a different object. Malsburg has explored this phenomenon from a modeling perspective to look at sensory segmentation with coupled neural oscillators (Malsburg and Buhmann, 1992).

Along these lines then, we explored oscillatory synchrony with a simplified visual binding question about image segmentation. We were presented with the following problem: Suppose you have as input a binary two dimensional array, which could be considered a static, black-and-white visual 'image.' Suppose this image was presented to a corresponding array of mathematically modeled neural oscillators. Motivated by the binding problem then, we want neurons which receive input from one 'object' (defined as a contiguous group of nonzero entries in our input array) to oscillate in-phase, while neurons receiving input from another 'object' also oscillate together, but out of phase with all other 'objects.' The question we set out to investigate then, was as follows: *what type of connectivity among these neurons is necessary to produce this specific oscillatory behavior given any binary input array?* In other words: how should we 'wire' the neurons, and with what type of connection (excitatory or inhibitory) should we have, in order to make the neurons corresponding to the same object to fire in-phase, but make the neurons corresponding to different objects to fire out of phase?

Having an idea of the neurobiology gives us an idea of what assumptions we are making in our proposed question about image segmentation.

For example, our largest simplification assumes a one-to-one correspondence between input photons and neurons in V1, even though this relationship is demonstrated only in the earliest level of processing at the retina. We further simplify the neurobiological problem of visual image segmentation by considering a flat input as opposed to a stereoscopic input. This input is binary and static, corresponding to a black and white still image with no overlapping objects, as opposed to visual input filled with moving multi-colored and textured objects. Finally, we assume there is a background to our image and propose that neurons corresponding to the background should not fire, whereas in reality identifying the background is an esoteric question at best.

The Mathematical Model

Hodgkin-Huxley equations

The mathematical model of our work is based on the reduced Hodgkin-Huxley system of equations, which contains three ODEs. See (Terman, 2003) for a thorough explanation of these equations:

$$v' = f(v, n) + I_{syn} + I_{app}$$

$$n' = \varphi \cdot g(v, n)$$

$$s' = \alpha \cdot (1 - s) \cdot s_{\infty} - \beta \cdot s$$

where

$$f(v, n) = -g_l(v - v_l) - g_k n^4 (v - v_k) - g_{n_a} m_{\infty}^3 (0.5 - n)(v - v_{n_a})$$

$$g(v, n) = \frac{(n_{\infty} - n)}{\tau_n(v)}$$

$$m_{\infty}(v) = \frac{1}{1 + e^{-(v + \theta_m)/s_m}}$$

$$n_{\infty}(v) = \frac{1}{1 + e^{(v - \theta_n)/s_n}}$$

$$\tau_n(v) = \tau_0 + \frac{\tau_1}{1 + e^{(v - \theta_n)/\sigma_n}}$$

$$s_{\infty}(v) = \frac{1}{1 + e^{-(v - \theta_s)/k_s}}$$

Note that the term I_{syn} in the first equation is very important. It describes the synaptic connections between cells. There are two basic types of synaptic connection: inhibitory and excitatory. Let's look at I_{syn} carefully. It has the form:

$$I_{syn} = s \cdot g_{syn} \cdot (v - v_{syn})$$

where s is called *synaptic strength*, and describes the releasing speed of neurotransmitters of the pre-synaptic cell. The parameter g_{syn} is called *synaptic conductance*, and describes the receptivity of the post-synaptic cell. These two terms combined describe the effectiveness of the synapse. The parameter v_{syn} dictates whether the synapse is excitatory or inhibitory.

Synchrony terms

A cell can have both excitatory and inhibitory synapses from other cells. So, I_{syn} is actually composed of two parts: I_{syn_e} and I_{syn_i} , which represent excitatory and inhibitory synapses respectively.

We know that if two excitatory cells are coupled, they will tend to fire together, or they will go “in sync”. Intuitively, one would think that two inhibitory-coupled cells will tend not to fire together, i.e., will go “out of sync”. This is true, and can be easily checked using computer simulations. If the two inhibitory-coupled cells are identical to each other and have the same initial state, then, during the computer simulation, it might take a very long time for them to go out of sync. But, if noise is introduced into the system, the two cells go out of sync after very few oscillations.

Modeling the connectivity

The first approach we took to solve this problem was based on the idea of proximal excitation and distal inhibition. Consider, if two neighbor neurons receive input from the same ‘object’ they will both depolarize. Intuitively, if they are coupled in an excitatory way, then they will tend to synchronize. Also, if they send inhibitory connections to more distal neurons, these are less likely to correspond to the same ‘object’ and so will prevent them from firing in phase. With this approach, we expected that even if there were some inhibitory connections within the same object, the stronger excitatory connections would override any desynchronizing effects.

To implement this idea of proximal excitation and distal inhibition, we used what we called a ‘Mexican hat’ approach (see Figure 1). Here, each cell was connected with strong excitatory connections to its nearest neighbors, shown by the red bars in Figure 1, while the next neighbors over were connected to cell n with inhibitory synapses, shown by the blue bars in Figure 1.

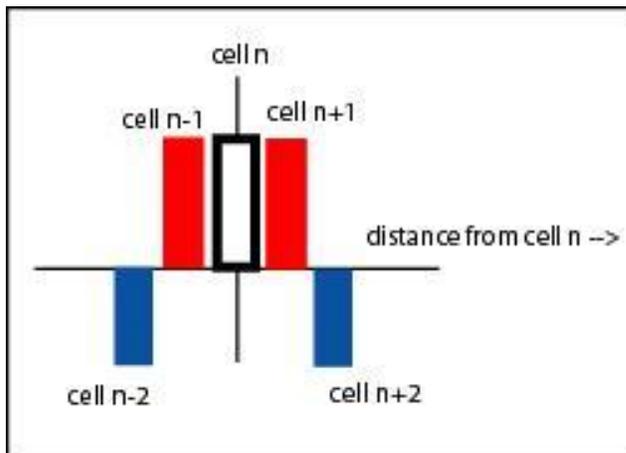


Figure 1. Mexican hat connectivity.

The edge effect

One problem quickly noticed is that the neurons which receive inputs from the edges of the input array do not have as many neighbors. To handle this asymmetry, we added self-loops as need to ensure a equal excitatory and inhibitory inputs were available to all neurons. This is illustrated in Figure 2.

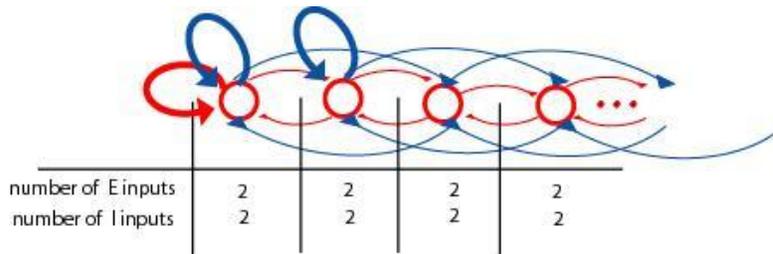


Figure 2. Edge self-loops.

Problems with the ‘Mexican hat’ connectivity.

Though this seemed to show some promise, we came to realize that there were several problems with modeling the neuronal connectivity in this way.

First, there is the problem of scope: how local is “local” and how distal is “distal”? Though this approach allowed flexibility in determining the range and strength of each connection (by, for example, making the nearest three neurons coupled in an excitatory way, and the next nearest five neurons coupled in an weakly inhibitory way) this is not enough solve the scope problem.

The concern was that this type of connectivity would not be able to generalize across all possible input arrays of all different sizes. For example, if we are given two small objects located at the very periphery of a large array, then we would need to ensure that the inhibitory scope of a neuron from one object synapsed onto a neuron receiving input from the other object. After considering cases such as these, we realized that the ‘Mexican hat’ approach was inadequately generalizable.

Our second concern addresses issues arising from neurobiological considerations. Though there is evidence for some cells which have both inhibitory and excitatory synapses onto other cells, these are not very commonly observed. The majority of cells have either excitatory or inhibitory projections onto other cells, and not both. With these two considerations in mind, we turn to a different, and more successful approach.

The Global Inhibitor

In this approach, we again used the idea of local excitation, but turned to global inhibition rather than distal inhibition. This allowed us to address our issues of adequate scope from above. Here, in the array of input neurons, every cell is excitatory and has nearest neighbor connections. We called this array of excitatory neurons the sensory layer. In addition to projecting to their nearest neighbors, they also project to the global inhibitor in the control layer. The global inhibitor is a single inhibitory cell that receives excitatory input from every sensory and projects inhibitory output to every sensory cell. This connectivity can be seen below.

With this model we had three different types of synapses, giving us many parameters with which to find the right balance of excitation and inhibition.

THE BASIC MODEL

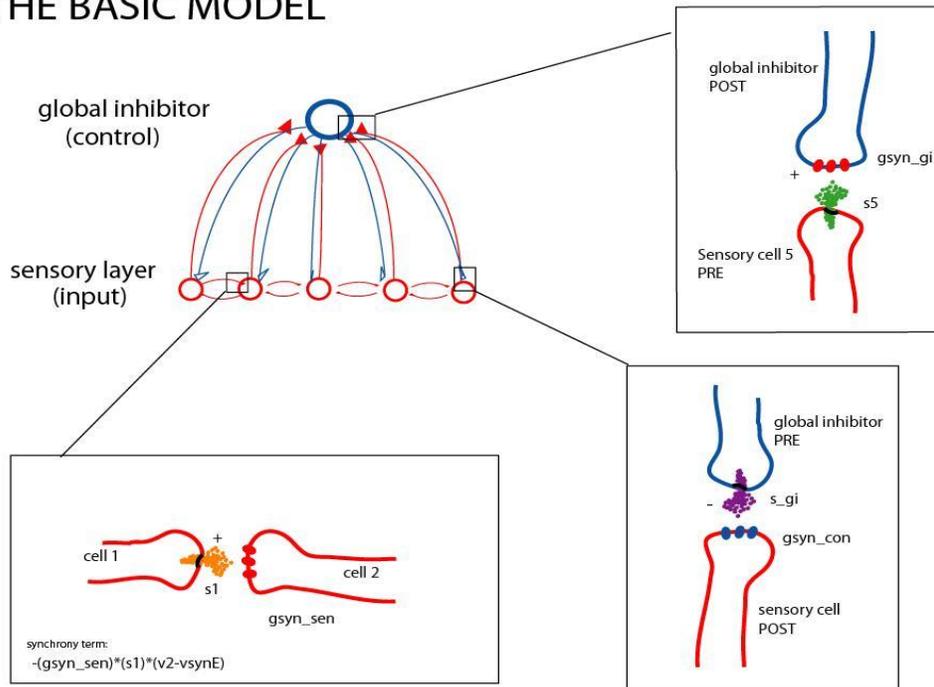


Figure 3. The control of the sensory layer using a global inhibitor.

Excitatory to excitatory. At the sensory level, each excitatory neuron can release neurotransmitter (where the amount and speed of release from cell n is reflected by ' sn ') and is receptive to neurotransmitters (where the conductivity of the synapse is represented by $gsyn_sen$).

Excitatory to Inhibitory. When a sensory cell synapses onto the global inhibitor, we assumed that it would release neurotransmitter in the same manner as it would onto its sensory neighbor. The global inhibitor has post-synaptic receptors whose conductance we modeled with the parameter `gsyn_gi`.

Inhibitory to Excitatory. Analogous to the above cases, we modeled the global inhibitor's neurotransmitter release with `s_gi`. The sensory layers receptivity to this controlling inhibitory signal was modeled with `gsyn_con`.

Given this neural architecture, we felt we had the capacity to model neuronal synchrony within objects and desynchrony across objects. Doing so involved exploring the parameter space and ultimately converging on a delicate balance between all the parameters involved. This was accomplished by keeping several heuristics in mind. (1) We needed enough excitation at the sensory layer to keep the objects in synchrony but (2) not too much or else neurons corresponding to the background would fire. Finally (3), we need to add in enough inhibition to create asynchrony between the object groups but (4) not so much as to desynchronize neurons corresponding to the same object.

The two-dimensional case

With this framework in mind, we then moved from a one dimensional input array to a two dimensional input array (and a corresponding two dimensional array of sensory neurons). To handle the increased complexity, we transitioned from modeling in XPP to MATLAB software. In Figure 3 we show the model for the 2-dimensional case (not all of the projections to and from the global inhibitor are shown).

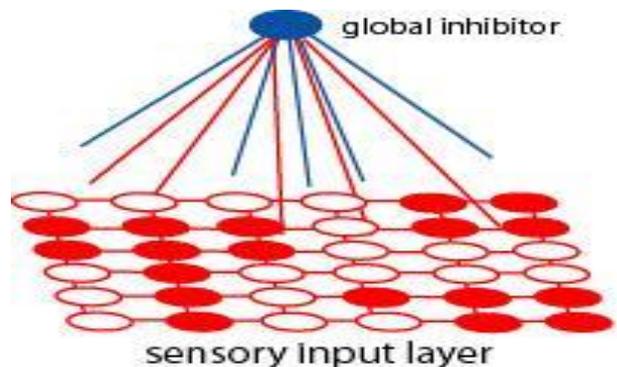


Figure 4. Two dimensional array of neurons and the global inhibitor.

Like the 1-dimensional model, the neurons in the 2-dimensional model are grouped into two layers. The edge effect was handled using self-connections, as before. More exactly, neurons receiving input from the corners of the arrays have two nearest neighbor connections and two excitatory self-loops. Similarly, edge-neurons have one added self-connection.

In order to see more clearly how the neurons fire according to different sensory inputs, we modeled such a network of neurons using MATLAB. Here, we used a 10 by 10 matrix of neurons. We tried different sensory inputs designed specifically to test the considerations described above:

- 1) 'Donut shape' (a square with a hole in the center). In this case, though the neuron that corresponded to the hole received many excitatory inputs from its sensory neighbors, it receives enough inhibition from the control layer so that it doesn't fire.
- 2) The letter 'B', which has two holes. As before, here neurons corresponding to both holes keep silent while other neurons receiving inputs fire in synchrony.
- 3) Multiple objects. For example, for three objects, we get the results shown below. The three objects rapidly go out of phase. Some of the neurons that do not receive sensory input tend to fire in the beginning, but are suppressed as the other objects come into synchrony, due to the effect that the global inhibitor.

Limitations of the model

If the number of objects in the image is too large (e.g., larger than five), then the model might not be able to differentiate between the objects by the phase of their corresponding neurons. Figure 5 shows such results for two neurons corresponding to different objects over a long period of time. As can be seen, the two neurons go in-phase, then out-of-phase, then in-phase again, and so on. On the other hand, this suggests that the neurons could also distinguish different objects by 'frequency coding'.

Further Work and Discussion

Moving objects

After being able to simulate a visual segmentation of fixed objects by using neural oscillators, we try now to extend our ideas for the case of moving objects in the visual field.

To do that, we basically consider a moving input function. We implement this idea using a one-dimensional array of 25 coupled neurons and a global inhibitor with the same circuitry as before. As we consider an object of size N (in the simulations we took $N=5$), say from left to right in our one-dimensional visual field, the input changes as in Figure 6.

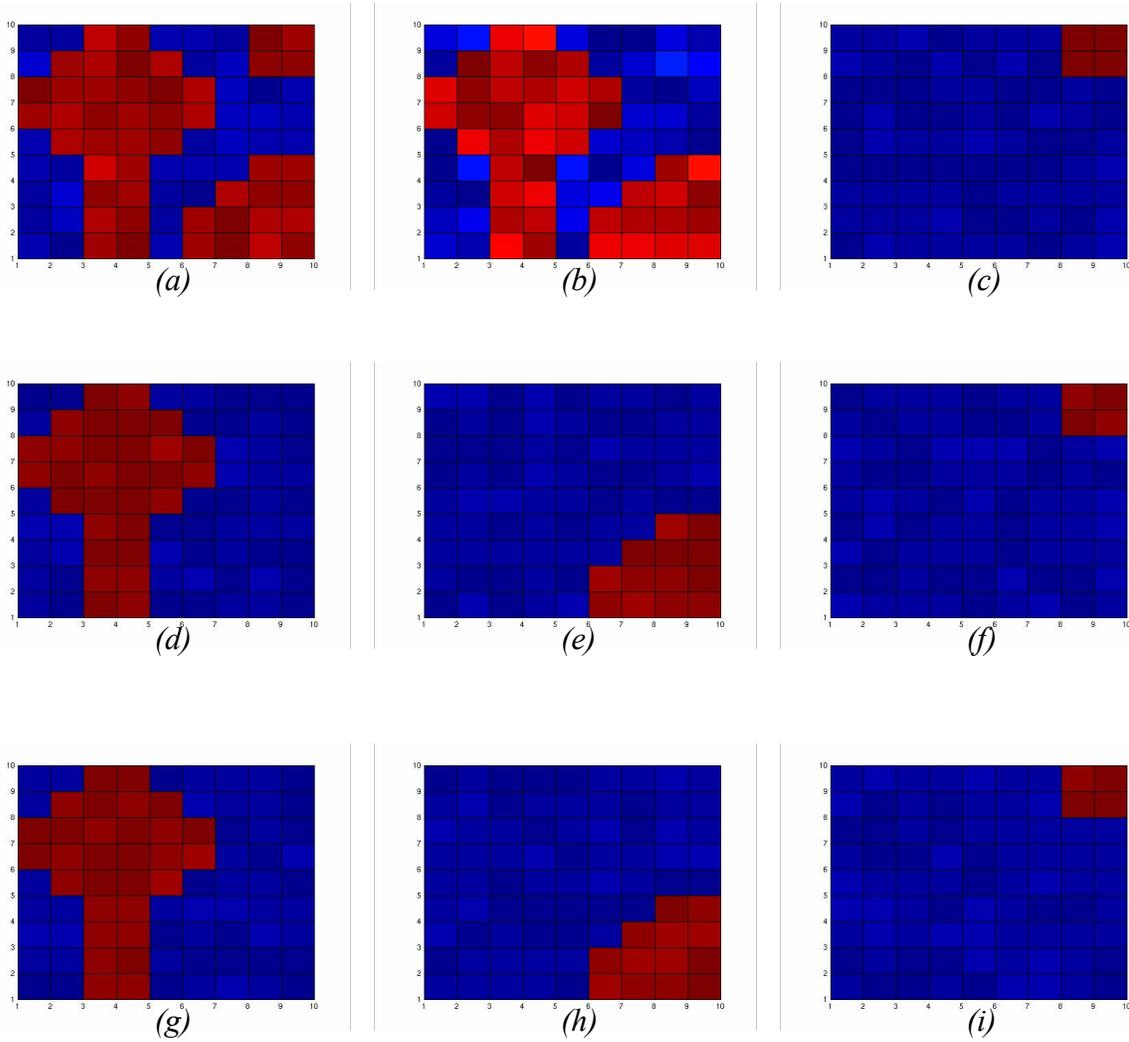


Figure 5. Three different objects in a two-dimensional network of neurons. Initially, all excited neurons fire in phase, see (a); after a short time, neurons corresponding to different objects go out of phase, while neurons corresponding to the same object remain in phase, see (d)-(i).

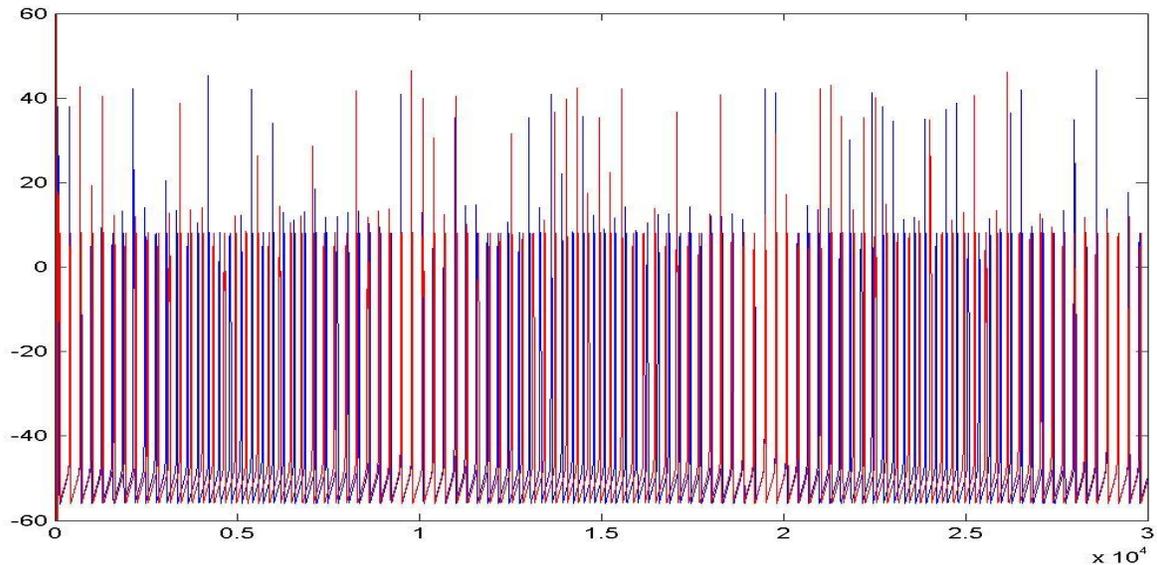


Figure 6. If the number of different objects is too large then neurons corresponding to different objects may not go out of phase with each other.

The results of the simulations in this case are quite promising. One problem we found was that the neuron representing the very front of the object would not fire in synchrony right away with the rest of the neurons. Rather, it showed irregular firing activity before settling into synchrony. We hypothesize that this is due to the sudden onset of input this neuron receives when the object moves into its 'receptive field' (i.e. array location). One way to address this would be to consider a much smoother transition of input to the neurons, e.g., to considering Gaussian-type input functions. With further extensions of this sort, one might try to implement stable and moving objects together or having objects moving at different speeds.

Though we were able to model image segmentation using within-object synchrony and across-object asynchrony, there are many limitations of this approach. For example, the biological system is obviously very complex; and a static array with two-dimensional sensory layers is an overwhelmingly large simplification.

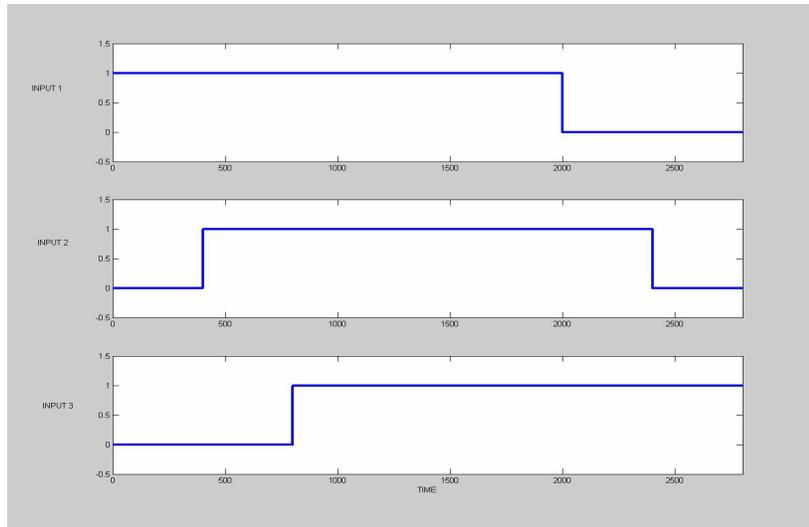


Figure 7. The moving input function.

Also, we took a more top-down approach to this problem, knowing we wanted to model object synchrony and looking at what neural architecture had the capacity to implement this. It would be interesting to see if our connectivity model is at all similar to neuronal circuitry at the sensory layers of the cortex or at other levels in the perceptual processing stream. Despite its limitations, our restricted analysis does suggest that this type of synchrony of oscillations could be a mechanism that the brain employs in its solution to the binding problem.

Acknowledgments. This work benefited from discussions with David Terman, and was supported by the National Science Foundation under Agreement No. 0112050.

REFERENCES

Eckhorn R., Bauer R., Jordan W., Brosch M., Kruse W., Munk M., Reitboeck H.J. (1988) *Coherent oscillations: a mechanism of feature linking in the visual cortex? Multiple electrode and correlation analyses in the cat*, Biological Cybernetics 60 (2) 121-130.

Kandel, E. R., Schwartz, J. H., Jessell, T. M. (2000) *Principles of Neural Science*. 4th Edition, McGraw-Hill, New York.

von der Malsburg C., Buhmann J. (1992) *Sensory segmentation with coupled neural oscillators*, Biological Cybernetics **67** 233-242.

von der Malsburg C., Schneider, W. (1986) *A neural cocktail-party processor*, Biological Cybernetics **54** 29-40.

Terman D., Wang D. (1995) *Global competition and local cooperation in a network of neural oscillators*, Physica D **81** 148-176.

Terman D. (2003) *An introduction to dynamical systems and neuronal dynamics*, Lecture Notes, Mathematical Biosciences Institute Summer Program, July 14 – August 1, 2003.