

A NEURAL MODEL FOR THE DETECTION OF TEMPORAL STRUCTURE

Dražen Domijan

*Department of Psychology, Faculty of Arts and Sciences
University of Rijeka, Slavka Krautzeka bb, 51000 Rijeka, Croatia
<ddomijan@ffr.hr>*

Abstract

According to the temporal correlation hypothesis, synchronization of neural activity in different spatial maps solves the feature binding problem. Here, a new model of visual working memory is proposed which is able to group synchronous temporal events without neural synchronization. Instead, the model is based on the difference in the firing rate. The model integrates discrete inputs over time and compares activity in different integrators. When the amplitude difference in integrators is large enough due to the different rates of evidence accumulation, temporal figure and background are distinguished in the working memory. Computer simulations showed that the model correctly groups events according to their deterministic or stochastic temporal structure. The model is robust with respect to the temporal noise and to the correlation between figure-ground events. Also, the model is able to explain visual prior entry and perceptual asynchrony between colour, motion and orientation.

Temporal correlation hypothesis (TCH) offers a solution to the problem of how visual cortex achieves binding of distributed features into unified precepts based on the timing of neural events (Singer & Gray, 1995). Neural signals that are grouped together in time are assumed to have a common cause and they are assumed to represent single object. On the other hand, neural signals that are decoupled in time correspond to different objects in the visual field. Besides neurophysiological investigations, it was suggested that if TCH is correct, its effect should be visible in psychophysical tasks also. Early investigations support the conclusion that the visual system is sensitive to the temporal dimension of the stimulation, but these findings were criticised on the ground of confounding spatial and temporal cues for grouping (Blake & Lee, 2005).

Recently, Guttman, Gilroy and Blake (2007) proposed that visual system is sensitive to the stochastic temporal structure which is defined by random point processes. In several experiments they showed how figure-ground discrimination depends on the type of point process utilised for definition of figure and ground regions. If figure and ground are defined by independent point processes or if they are defined by the same point process but with large delay in the ground region compared with the figure region, figure-ground discrimination is successful. On the other hand, if figure and ground regions are defined by the same point process which is delayed in the ground region for small number of frames, discrimination performance declines.

In order to account for the psychophysical performance in the figure-ground discrimination tasks based on the temporal information I proposed a neural network model which uses difference in firing rates as a representational format. The model assumes receiving analogue input and producing output that will discriminate figure from ground by different amplitudes of neural activity.

Method

Here, I proposed a new neural network model of visual working memory that is capable of grouping temporal signals according to their phase lag (deterministic structure) or to their accumulated phase lag (stochastic structure). The model is based on the assumption that visual system is segregated into two processing streams: 1) fast system dedicated to the detection of transient changes in visual field, and 2) slow system which integrate incoming signals over time and exhibits sustained response to the input. Both systems provide input to the model working memory (WM) network which is able to discriminate signals arriving at different times or with different amplitudes. Transient system is modelled as a change detector which produces short pulses of activity at the moment of change in the visual field. Sustained system is modelled as a neural integrator where the network activity is accumulated over time. Activity accumulation depends on the number of transient signals that simultaneously arise in the visual field. Accumulation is possible only when the number of transient signals is small compared to the total number of stimuli which corresponds to the figure region.

Transient and sustained systems feed their output to the working memory network. It is modelled as a recurrent neural network with lateral inhibition and dendritic inhibition (Domijan, 2003). Lateral inhibition implements competition between spatial locations. In order to distinguish signals arriving at different times another type of inhibition is introduced which operates on the dendrites of the network nodes. Dendritic inhibition serves as protective gate that modulates the impact of lateral inhibition on the target node. If dendritic inhibition is stronger than lateral inhibition it will completely isolate the target node from lateral inhibitory influences. Due to the self-excitation, the target node will reach maximal level of activity and it will represent figure. On the other hand, if dendritic inhibition is weaker than lateral inhibition, the target node will not be able to attain maximal activity level and it will represent background.

Mathematically, the network nodes are modeled with non-linear shunting equation (Domijan, 2003):

$$\frac{dx_i}{dt} = -Ax_i + (B - x_i)(I_i^T(t) + I_i^S(t) + f(x_i)) - (C + x_i) \sum_{j \neq i} g[f(x_j) - f(x_i) - T], \quad (1)$$

Where, x_i , is activity amplitude (i.e., mean firing rate) of the node at spatial position, i . Term, A , is a passive decay which drives activity toward zero if there is no input; B (C) defines the excitatory (inhibitory) saturation point, that is, upper (lower) boundary for activity level that could be obtained. Function, $f(z) = z$ if $z > 0$ and $f(z) = 0$ if $z \leq 0$, describes half-wave rectification commonly used as an output for the node's activity. Function, $g[z] = 1$ if $z > 0$ and $g[z] = 0$ if $z \leq 0$, describes non-linear dendritic computation and, T , is a threshold for dendritic activation. T controls the amount of difference between the target node activity and lateral inhibitory signals. Lateral inhibitory signals, $f(x_j)$, are modulated by dendritic inhibition, $f(x_i)$, before they can influence target cell, x_i . For simplicity, inhibitory interneurons which mediate inhibitory interactions are not explicitly represented in the model but are incorporated into the eqn (1).

Transient input, $I_i^T(t)$, is set to, 1 , when the change occur at time, t , and to, 0 , at all other times. Sustained input is given by the

$$I_i^S(t) = f(y_i - Tr) \quad (2)$$

Where, y_i , describes the activity of the neural integrator which accumulates transient inputs as in

$$\frac{dy_i}{dt} = I_i^T(t) - y_i \sum_j w_j I_j^T(t) / N. \quad (3)$$

Term, Tr , is a threshold for activation of the working memory network, and, $f()$, is an output function described above. Neural integrator counts the number of occurrences of the transient input. However, such activity accumulation is counteracted by the global inhibitory signal which is proportional to the ratio of the number of transient signals, I_j , and total number of stimuli present in the visual field, N . Global inhibition is modulated by the synaptic weights, w_j . Inhibition enables integrator to count only those transient signals that arise from the small region of space (figure). On the other hand, transient signals arising from the large region (background) or signals that simultaneously arrive from the figure and background regions will cancel excitatory input to the integrator. Therefore, integrator is sensitive only to those transients that occur isolated in time and space.

Strength of the neural activity in the WM network depends on the time of input arrival or on the amplitude of the simultaneously arriving inputs. Nodes which receive input earlier in time will start to increase their activity earlier, so lateral inhibition, arriving later in time, will not be able to influence these nodes. On the other hand, all the nodes, that simultaneously receive input signals of equal amplitude, will not inhibit each other due to fact that they all have the same activity level and dendritic inhibition will cancel out all lateral inhibition between them. However, if inputs arrive with different amplitudes, the network will prefer nodes receiving stronger input. Therefore, temporal discrimination in the model arises from asymmetric lateral inhibition which favours nodes with stronger activity over nodes with lower activity levels or nodes which are stimulated earlier over nodes which are stimulated later in time.

Results

Computer simulations are performed by numerical integration of equations 1 and 3 using Runge-Kutta method. Network parameters were set as follows: $A=1$; $B=2$; $C=0$; $T=0$; $Tr=1$; $w_j=2$ for all j . Temporal evolution of network activity is plotted as a function of spatial position in the network. Amplitude of the neural activity is marked with the shades of gray with white indicating maximal activity and black indicating no activity. Input to the network is a series of short pulses that arises at the moment of change in the environment.

Figure 1 shows the network response to the temporal pattern with deterministic structure. In other words, figure is distinguished from background by constant temporal difference in input onsets for figure and background regions. The network is able to detect this temporal difference at the beginning of the stimulation and continue to actively maintain representation of the figure region by sustained neural activity during the whole trial (Fig 1a). Therefore, later input signals will not have any impact on the neural activity. It is sufficient that the temporal difference exists at the beginning of the trial. This is consistent with the finding of Beaudot (2002) who showed that initial phase lag is sufficient to discriminate figure from ground.

When there is no temporal difference between figure and ground, the model treats the whole visual field as a single group (Fig 1b). However, if input from the figure region is enhanced as when attention is directed to it, figure region is again visible despite the lack of temporal difference of input arrivals (Fig 1c). In this way, the model simulates visual prior

entry, i.e., precedence in processing of attended stimuli (Shore, Spence, & Klein, 2001). Interpretation of visual prior entry arising from this simulation is that the same neural network which is sensitive to the temporal difference of input is also sensitive to the amplitude of the input signals. Therefore, stronger signals which arrive simultaneously with the weaker signals will be perceived as arriving earlier in time.

It should be noted that the network response in figure 1b does not imply that input synchrony must be perfect in order to detect simultaneity. How much difference is tolerated depends on the threshold for dendritic inhibition, T . If threshold is set to zero, then perfect synchrony is needed to group all input signals into a unified percept. However, if threshold is set to some non-negative value, small difference at the start of activity build-up will not disrupt temporal grouping.

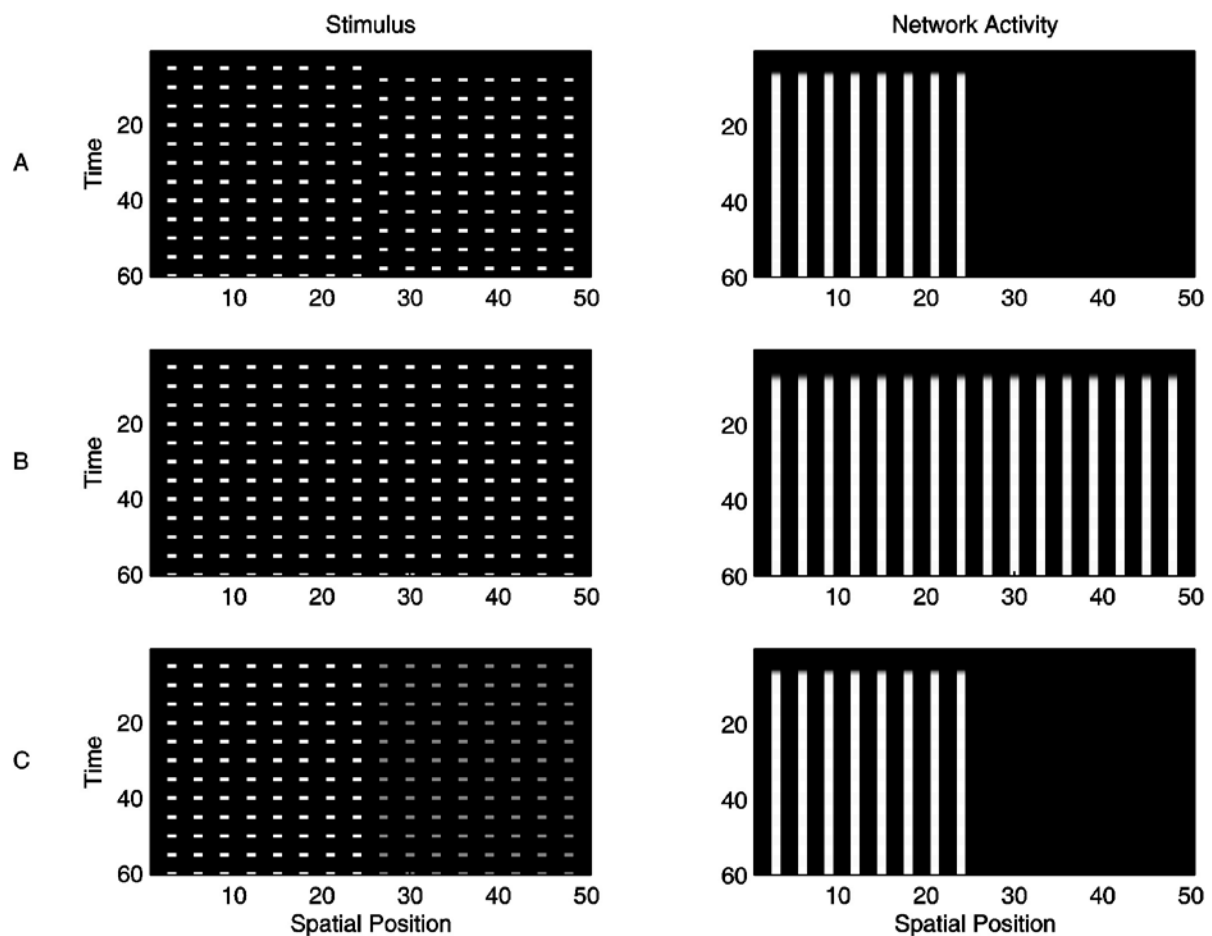


Fig. 1. Detection of deterministic temporal structure.

Figure 2 illustrates the network output for input with stochastic temporal structure. Here, I simulated the Experiment 3 of Guttman *et al.* (2007) and used the same point processes. When figure and ground are defined by independent stochastic point processes, figure is easily discriminated. The reason for this is a difference in the magnitude of activity in the sustained system. Transient system could not discriminate such situation because both figure and ground point processes start at the same time. However, in the sustained system, activity builds up according to the density of input signals within figure and ground regions. Within both regions there are periods of time when more signals arrive at one of the regions. Such differences are sufficient to produce different activity amplitudes in the sustained system for

the figure and ground area. Activity magnitude for neural integrators in the figure region is sufficient to reach threshold and to trigger activity in the WM network.

In the case when figure and ground are defined by the same point process which is delayed in the ground region for four frames, observers had difficulty in finding figure. The same problem is evident in the WM network response (Fig 2b). Here, the sustained system attempts to integrate signals over time but the density of signals is not sufficient for activity to reach threshold for working memory network. Therefore, the WM network does not show any activity and it is not able to discriminate figure from ground. When two point processes had the same temporal structure but the ground region is delayed for eight frames with respect to the figure, the discrimination performance returned to the level as if the figure and ground are independent processes. The WM network behaves in the same way. The sustained system is again able to reach threshold for the WM network due to the fact that longer delays creates greater density of input signals within certain time windows (Fig 2c).

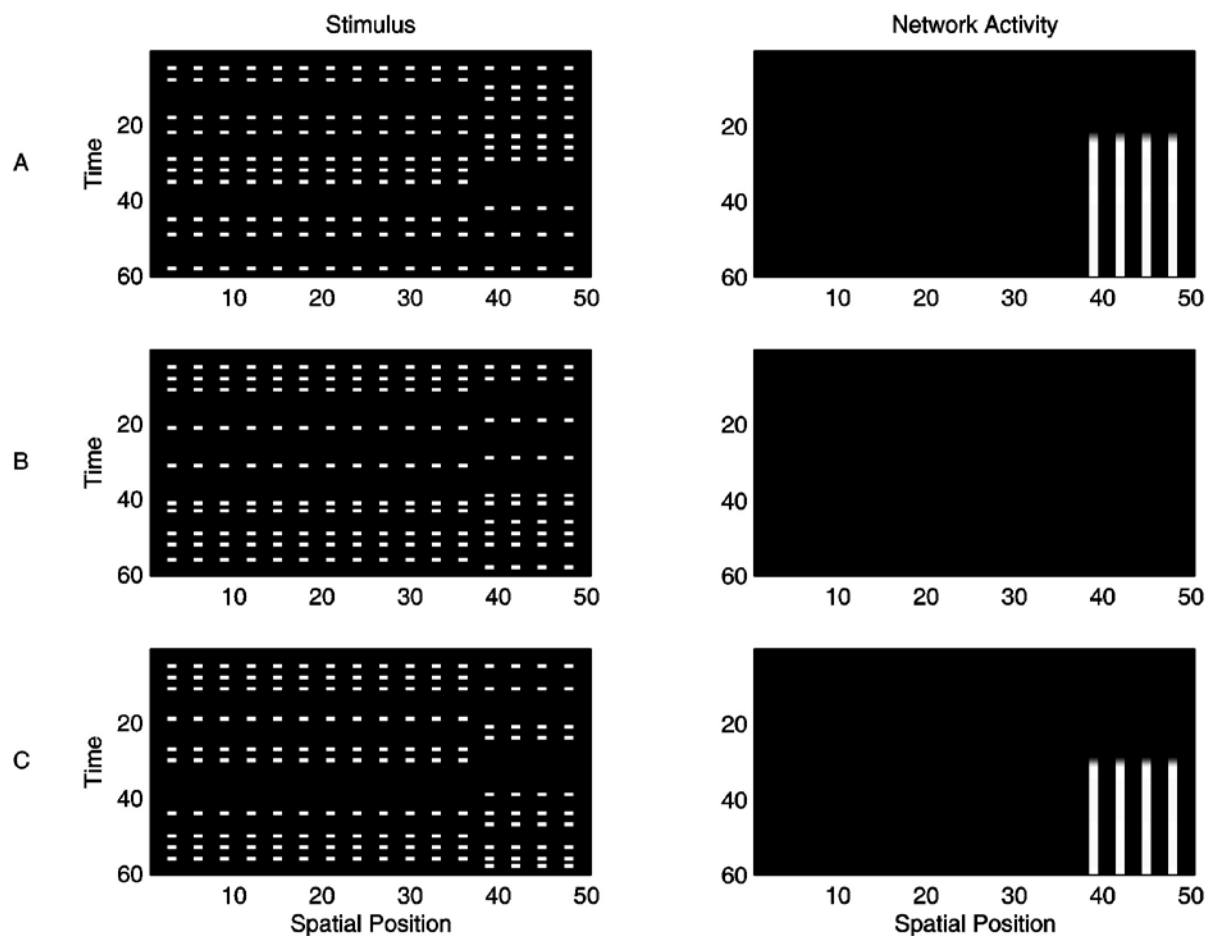


Fig. 2. Detection of stochastic temporal structure.

In a similar vein the model is able to explain the results of Experiment 4 of Guttman *et al.* (2007). In this experiment figure region is seen when it is defined by the jittered point process (ground-figure-figure-ground), but it is not seen when regular exchange between figure and ground occur (figure-ground-figure-ground). In the first case, the model is able to accumulate activity in the figure region due to the double change in it at the time window when there is no change in the ground region. In the second case, global inhibition disables activity accumulation and prevents figure-ground separation.

Discussion

Although the psychophysical research on the role of temporal grouping in vision has been motivated by the goal to find support for the temporal correlation hypothesis, here it was shown that all the available evidence could be simulated using the neural network model which is sensitive to the amplitude of the input signal and which produce different activity amplitudes for figure and ground region. The model uses two distinct types of input signals: transient and sustained. Transient signals contribute to the detection of deterministic temporal structure by providing different onset time for activity build-up in the WM network. Sustained signals contribute to the detection of stochastic temporal structure due to the integration of several transient signals over time. Therefore, sustained signals carry information about density of transient signals over certain time window. Different density of transient signals indicates difference in stochastic point processes and could reveal spatial structure of figure.

The distinction between transient and sustained systems and their contribution to the detection of temporal structure has been used previously in order to explain properties of perceptual asynchrony. It refers to the striking inability to perceive simultaneous changes in feature dimensions such as colour, motion or orientation. For instance, colour change appears to precede motion direction changes for about 80 ms (Bedell *et al.*, 2003). However, perceptual asynchrony is greatly reduced when temporal order judgment task is used. Task sensitivity points to the fact that different tasks may utilise transient and sustained system to a different degree. Present model is consistent with such interpretation and offer further support for a two-system account of the temporal processing in the visual system.

Acknowledgments

This work was supported by the Bial Foundation grant 80/06 and by the Croatian Ministry of Science, Education and Sport grant 009-0362214-0818.

References

- Beaudot, W.H.A. (2002). Role of onset asynchrony in contour integration. *Vision Research*, 42, 1-9.
- Bedell, H.E., Chung, S.T.L., Ogmen, H., & Petel, S.S. (2003). Color and motion: which is the tortoise and which is the hare? *Vision Research*, 43, 2403-2412.
- Blake, R., & Lee, S.-H. (2005). The role of temporal structure in human vision. *Behavioral and Cognitive Neuroscience Reviews*, 4, 21-42.
- Domijan, D. (2003). A mathematical model of persistent neural activity in human prefrontal cortex for visual feature binding. *Neuroscience Letters*, 350, 89-92.
- Guttman, S.E., Gilroy, L.A., & Blake, R. (2007). Spatial grouping in human vision: Temporal structure trumps temporal synchrony. *Vision Research*, 47, 219-230.
- Shore, D.I., Spence, C., & Klein, R.M. (2001). Visual prior entry. *Psychological Science*, 12, 205-212.
- Singer, W., & Gray, C.M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annual Review of Neuroscience*, 18, 555-586.