Greg Amis
CN 530
29 November 2004
Simulation Assignment 3

Analysis of the Grossberg and Rudd Motion-BCS Model: Travelling "G-waves"

CN 530: Neural and Computational Models of Vision

29 November 2004

Abstract

Apparent motion is a phenomenon of visual perception in which smooth motion is perceived even when no smooth motion exists in the visual stimulus. The motion boundary contour system (BCS) model is briefly described as a biologically plausible model for apparent motion perception. A static BCS defines luminance boundaries within the image. These boundaries are converted into local motion signals that charge as an input is presented and decay when the input is removed. These local motion signals are then convolved with spatial Gaussian curves, in order for nodes with spatially separate inputs to impact each other's activity over a formalized time course. The maximum values of these motion signals over space at a given time represents the global motion signal and is a computational analog of the psychophysical percept. A 32-node MATLAB implementation for this model is presented and simulations are performed using a two-flash stimulus with a 19-node spatial offset and an inter-stimulus interval (ISI) of zero. We find that the model registers a smooth motion signal in response to this two-flash stimulus simulating the psychophysical phenomenon of apparent motion. The system achieves this without modeling motion as a purely temporal object. The motion BCS provides further support for the boundary contour system as a viable, unified model of visual processing.

1 Introduction

In this report, we will analyze the performance of a version of the motion boundary contour system (BCS) presented by Grossberg and Rudd (1989). This model attempts to simulate and explain the phenomenon of apparent motion, whereby motion is perceived even when none is actually occurring. By modeling apparent motion perception as an extension of Grossberg's static boundary contour system, this model serves to explain previously unmodeled phenomena and to provide further support to BCS as a unified theory of visual processing.

Apparent motion has been studied extensively in the context of using carefully timed flashes of two spatially proximal stimuli, and that is the type of scenario we will focus on here.

2 Methods

All plots are the results of implementing a modified, 32-node version of the model in MATLAB and presenting to the model the stimulus pattern in Figure 1.

The following modifications and assumptions were used when implementing the model, for simplicity, clarity, and consistency with the original Grossberg and Rudd (1989) article.

- As in the Grossberg and Rudd (1989) publication, the oriented contrast filters are assumed to respond at the center of input stimuli (e.g., $i = 3$ and $i = 24$, for the stimulus in Figure 1), rather than at the edges.

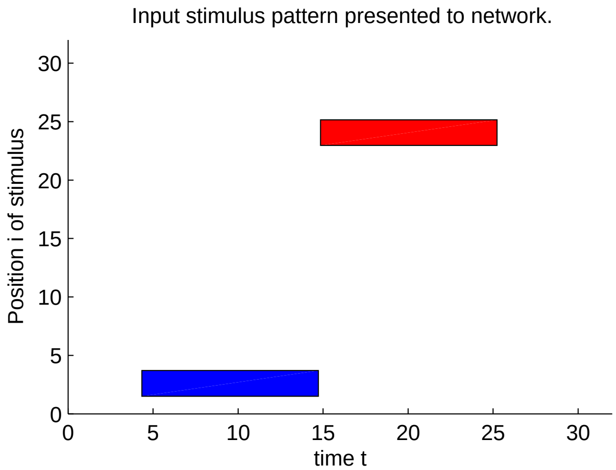

Figure 1: Input stimulus pattern presented to the network. An input stimulus of strength 1 was presented at nodes 2, 3, and 4 from time $t = 4$ to $t = 16$. A second input stimulus, also of strength 1, was presented at nodes 23, 24, and 25 from time $t = 16$ to $t = 28$. These temporal properties imply an inter-stimulus interval (ISI) of zero.

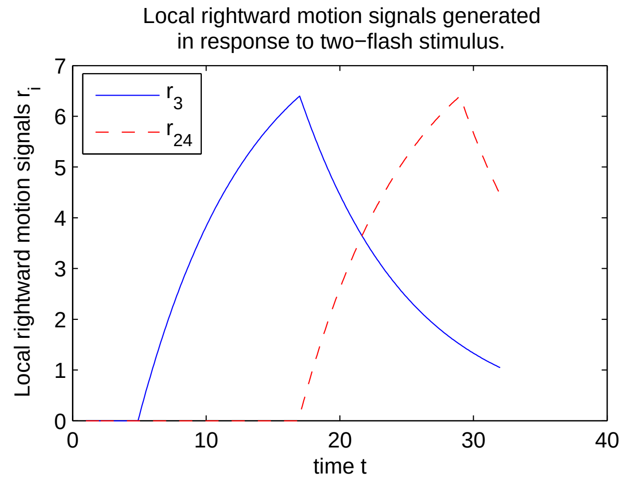

Figure 2: Local rightward motion signals generated by Levels 1-4 in response to a two-flash stimulus. Because of our assumptions in and simplifications of these initial levels, the local motion signals demonstrate the typical "RC curve" of a leaky integrator. Please see text for a detailed interpretation.

- The parameters $A = 0.12$, $B = 0$, $H = 1$, and $K = 12$ were used.
- Because we presume the system will detect a rightward motion, our implementation does not include the leftward motion components.
- The unoriented transient-response filters (Level 3) were not considered in our implementation. I.e., y_i^+ and y_i^- always equal 1.

All numerical integrations were produced via the forward Euler method using a time step of 0.1.

3 Results and Discussion

3.1 Local rightward motion signals

Figure 2 displays the local rightward motion signals (r_i) generated by Levels 1-4 at positions $i = 2$ and $i = 24$ in response to the input stimulus described in Figure 1.

When the stimulus is first presented to node $i = 3$ (at time $t = 4$), the response of r_3 "charges" in the same way that an electrical capacitor does: the rate of increase is initially very large but decreases linearly over time. When the stimulus is turned off (I_3 becomes zero at $t = 16$), r_3 decreases at a rate proportional to its value, in a fashion similar to an electrical capacitor discharging. A similar curve occurs at node $i = 24$ when I_{24} is presented (from time $t = 16$ to $t = 28$). Note that because r_{24} begins charging at the same time that I_3 turns off and before r_3 has had enough time to decay to zero, the two motion signals have significant temporal overlap. This overlap will prove to be critical in later levels of processing.

As was mentioned in the Methods section, we greatly simplified the first four levels of the model. With $B = 0$, the oriented sustained-response filters in Level 2 (x_{ik}) are no longer shunting but simple, additive, leaky integrators in the form $\frac{d}{dt}x_{ik} = -Ax_{ik} + J_{ik}$. And by assuming that the unoriented transient-response filters in Level 3 (y_i^+ and y_i^-) are always equal to one, the local rightward motion signals reduce to $r_i = x_{iL} + x_{iR}$. Since we are disregarding the leftward motion, $r_i = x_{iR}$, and therefore, $\frac{d}{dt}r_i = \frac{d}{dt}x_{iR} = -Ax_{iR} + J_{iR} = -Ar_i + J_{iR}$. And since we are having the oriented contrast filters respond to the center of stimuli, then $J_{iR} \simeq I_i$, so:

$$\frac{d}{dt} r_i = -Ar_i + I_i$$

This mathematical reduction explains our curves in Figure 2. When $I_i > 0$ and r_i is low, $\frac{d}{dt} r_i \simeq I_i$, causing the initial steep slope. As r_i increases, Ar_i increases, counteracting the impact of I_i on the rate of increase. When $Ar_i = I_i$, we reach an equilibrium point of $r_i = \frac{I_i}{A}$. In this case, that would be $r_3 = \frac{1}{0.12} = 8\frac{1}{3}$, however, that point is never reached, because I_3 is turned off before r_3 equilibrates. When I_3 is zero, $\frac{d}{dt} r_3 = -Ar_3$ and r_3 thus decays at a rate proportional to its value.

3.2 Trajectory of Maximum Motion Signal

When these local motion signals are convolved with a spatial Gaussian function, we begin to have some spatiotemporal interaction between the motion signals generated by the two stimuli.

Before we consider a plot of the simulation result, let's discuss what we would expect to happen to the motion signals.

If we had only one input stimulus, then the Gaussian wave produced by that stimulus would be centered at the same node as the stimulus center, and it would grow and shrink over time according to the local motion signal curve illustrated in Figure 2. Its maximum would always have the same spatial location.

Now let's consider if we added a second input stimulus with the same temporal onset and offset but different spatial position. In this case, the two Gaussian waves would grow and shrink in parallel over time. To the degree that the waves overlap in space, they would sum in positions between their centers. However, the maximum signal values would remain at the centers of each stimulus pattern and will never occur elsewhere.

Now that we have considered these two cases, let's return to the stimulus pattern that is the focus of this report, where the two spatially separate stimuli are temporally offset, with an inter-stimulus interval (ISI) of zero. Again, when the first stimulus is presented, the Gaussian curve grows, with its maximum at the node where the stimulus is centered ($i = 3$). When that first stimulus is turned off and the second stimulus is turned on, there is a point in time where the first train of Gaussians will be decaying from its maximum and the second train (generated due to the second stimulus) is still charging up. It is during this period that the summing of the two Gaussians at nodes in between the stimulus centers will be greater than the motion signal at the stimulus centers themselves.

Figure 3 illustrates this phenomena. At time $t = 16$, the local motion signal generated from the first stimulus is nearly full charged, and because no other stimuli exist at this time, the Gaussian curve's maximum is at the center of the stimulus ($i = 3$). At times $16 < t < 24$, the local motion signal generated from the second stimulus is beginning to charge while the local motion signal from the first stimulus is quickly decaying. As we predicted in the preceding paragraph, the summing of the two Gaussians in between the stimulus centers ($3 < i < 24$) is greater then the signal at the stimulus centers. As time t approaches 28, the local motion signal from the first stimulus is nearly fully decayed and the local motion signal from the second stimulus is nearly fully charged, so the sum of the two Gaussians reaches its maximum near the center of the second stimulus ($i = 24$).

Figure 4 plots the position of the maximum Gaussian-convolved motion signal value as a function of time, giving us a more distinct picture of this temporal shift of the maximum. Even though the stimuli only flash and do not move and even though the Gaussian waves generated by these stimuli do not spatially shift over time, the maximum motion signal value moves in a smooth, sigmoid trajectory from the first stimulus center to the second.

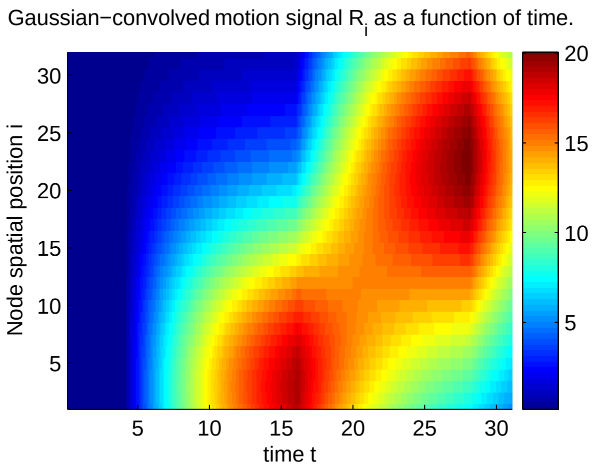

Figure 3: Plot of Gaussian-convolved motion signals R_i over time and space. The color of a position represents the value of $R_i(t)$. As indicated by the color bar, red areas indicate a larger value for $R_i(t)$ and blue areas indicate a smaller value. Please see text for details and interpretation.

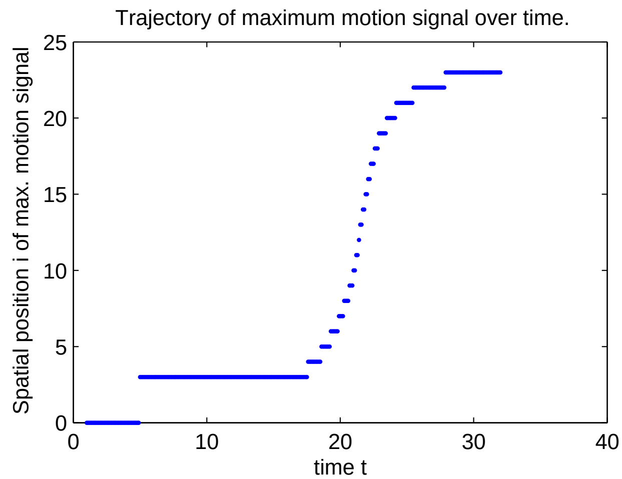

Figure 4: Plot of the position i of the Gaussian-convolved motion signal maximum as a function of time. See text for details and interpretation.

3.3 van Santen and Sperling "temporal object"

In 1984, J.P.H. van Santen and G. Sperling published an article in the *Journal of the Optical Society of America A*[2], postulating that motion (especially apparent motion) should not be modeled in terms of a spatial object occupying different locations at different times. Rather, they believed that one should "think of motion as involving a temporal object (luminance modulation pattern that occurs at different points in time at different locations)..."[2].

Does the Grossberg and Rudd model fully follow this advice? If so, each node (each spatial position) would have the same fluctuation of motion signal over time, with the exception of a shift forward or backward in time. To determine this, let's reexamine Figure 3. If we take slices of this figure parallel to the temporal axis, then each slice represents the change in motion signal strength for each spatial node over time. If each of these slices were congruent up to a temporal shift, at the very least, they would each have the same maximum value. At first glance, this does not appear to be true. Around node $i = 21$, we never achieve quite the same "red" values that we do at nodes $i = 3$ and $i = 24$, regardless of a temporal shift.

If we plot these three slices (as in Figure 5), we see that the Gaussian-convolved motion signals (R_i) varies differently over time for each node. They are not just the same pattern shifted in time. Therefore, the Grossberg and Rudd model does not fully follow the advice of van Santen and Sperling in that their representation of motion is not solely a "temporal object", invariant in space.

4 Conclusion

We have investigated the performance of a simplified version of the Grossberg and Rudd motion boundary contour system. In response to a two-flash stimulus with an inter-stimulus interval of zero, the model detects a smooth motion signal even though no motion occurs in the stimulus. This response simulates the psychophysical phenomena of apparent motion, where a visual stimulus devoid of motion can elicit the perception of motion.

While this model does not take the advice of van Santen and Sperling in modeling motion perception in terms of a purely temporal object, it does ad-

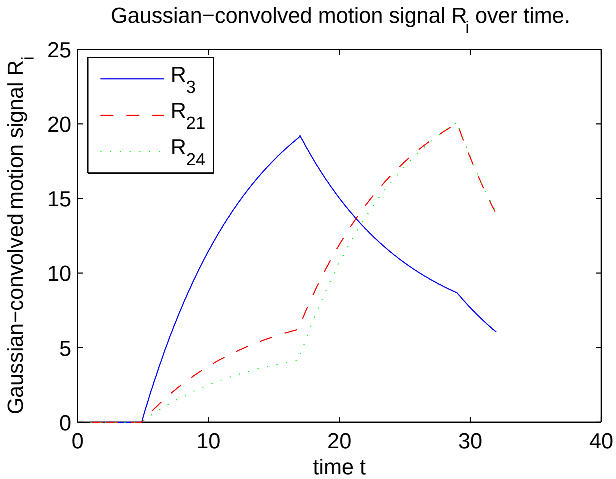

Figure 5: Plot of Gaussian-convolved motion signals (R_i) over time for three specific spatial nodes, $i = 3$, $i = 21$, and $i = 24$. Note that each node's profile is distinct and not just a temporally shifted version of the others.

equately explain the phenomena of apparent motion. Furthermore, the fact that this model was constructed via an extraordinarily simple addition to the static BCS model provides further support for the BCS as a viable, unified model of visual processing.

References

[1] Grossberg, S. and Rudd, M. E. (1989). A neural architecture for visual motion perception: Group and element apparent motion. *Neural Networks*, 2, 421-450.

[2] van Santen, J. P. H. and Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America A*, 2(2), 300-321.