Emergence of Border & Surface Completion (both Spatial and Temporal) in a Flowcentric Model of Narrow Slit Viewing

David Pierre Leibovitz (dpleibovitz@ieee.org) &
Robert L. West (robert_west@carleton.ca)
Institute of Cognitive Science, 2201 DT, Carleton University
1125 Colonel By Drive, Ottawa, Ontario, K1S 5B6 Canada

Abstract

In this paper, we describe a model of narrow slit viewing that deals with both spatial and temporal completion for borders and surfaces. The model is based on functionality derived from the dynamic interactions of a neural model. We contrast this model with FACADE, which models vision using neural models of modules corresponding to functionality.

Keywords: Border completion; Emergence; Emergic Cognitive Model; FACADE; Flowcentric; Slit Viewing; Spatial; Surface completion; Temporal; Unified Modeling.

Introduction

With regard to border and surface completion, the complexity of FACADE (e.g., Grossberg & Rudd, 1992) will be contrasted to the simplicity of ECM (Leibovitz, 2013a). However, it must be stressed that this paper is less about comparing these models and more about opening up epistemic options for analysis and decomposition as well as in explicating unrecognized modeling tradeoffs. To ground the discussion, we will mostly focus on demonstrating and explaining the emergence of border and surface completion effects within a flowcentric model of narrow slit viewing (Leibovitz, 2013b).

FACADE exemplifies a class of connectionist models with large-grained systems engineered to directly and isomorphically realize large-grained functions such as border and surface completion under motion. Although, the model is still being refined (Grossberg, Léveillé, & Versace, 2011; Grossberg, 2010), it represents a style of modeling and analysis prevalent throughout cognitive science and neurobiology. In contrast, ECM uses the interactions between finer-grained, but more complex units to create emergent functions capable of producing the same phenomena. As such, ECM can be thought of representing a dynamic systems approached based on a lower level of analysis. The goal of the ECM approach is to generate high level functional complexity as an emergent property of relatively simple, lower level interactions

FACADE

FACADE class models are quite complex in terms of the number of unique neural units, subsystems and parameters. For example, FACADE includes a *static* boundary contour system (BCS), a feature contour system (FCS), as well as a separate *motion* BCS. An example application of FACADE required 6 high-level kinds of neural units or subsystems involving 10 different interactions with parameters governed under 36 equations (Hu, Zhou, & Wang, 2011). However, in general, these solutions are not robust and highly sensitive to parameter changes. A more detailed model lists 62 parameters under the control of 43 equations (Berzhanskaya, Grossberg, & Mingolla, 2007). These models use low-level neural circuitry to produce distinct higher-level functions that can be thought of as distinct modules that interact.

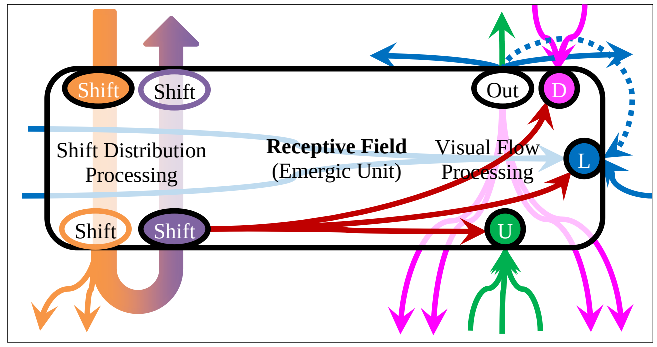


Figure 1: Functional Interactions in an Emergic RF

ECM

ECM is also a connectionist model but decomposes behaviour into finer *functional* grains that correspond to recurrent interactions of bottom-up, top-down and lateral flows of information. ECM has a single Receptive Field (RF) unit that is the locus of two functions and determines their interaction (see Figure 1). The RFs are arranged in a spatiotemporal hierarchy as indicated in Figure 2 (and Figure 7), but the behaviour of ECM can be fully understood by focusing solely on the RFs as we do in this paper.

On the visual stimuli processing side within each RF (the right hand side), the bottom-up, top-down and lateral flows of information interact according to the *handling missing data* function. This will be responsible for spatial completion according to the following algorithm: if a bottom-up visual value is missing (due to an eye blink or blind spot) then use a corresponding lateral value, and if that value is also missing, use a corresponding top-down value. Then send the value back out. The single output port directs the value up, down and laterally with additional local fan-out to allow for visual content to "shift". While the content of these three flows is similar, their usage differs. For example, lateral flows have the same spatial resolution as their RFs and produce memory-like effects. Spatiotemporal summation occurs naturally at all input ports, and so the RF has no further computation to perform.

Note that when visual information first enters the system at the Photoreceptor units, it is tagged by its spatiotopic extent within structured *emergic values*. This spatial extent is indicated within Figure 7. The bottom-left most RF has received a blue stimulus value with intensity of 20±0.0 and a spatial extent (on the x-coordinate) of 1±1.0. As the visual information flows up, the spatial extent increases but is narrower than expected. This is because information is summed according to statistical sampling techniques. The spatial extent is actually indicated by a single standard deviation from the mean.

On the left hand side of each RF, motion compensation or shift information for advanced eye motor plans is distributed down the RF hierarchy and reflected back up as part of the *maintaining information coherence* function. This allows the RFs to maintain their *logical* spatiotopic extents under eye motion. This is arranged such that when visual information flows upwards, it is synchronised with actual motor plans for moving the eyes (i.e., saccades). The input ports named U, D and L (having an updated spatiotopic extent) only accept visual values tagged with the same extent via dendritic processing (Leibovitz & West, 2013). Technically, each port sums up the input values weighted by the percentage of spatial overlap between the spatial extent tagged along with the visual values, and the spatial extent of the RF.

In this way, the two functions interact within the locus of every RF, and this is indicated via the red coloured lines within the RF of Figure 1. Although it appears as if the visual values shift, their tagged coordinates remain spatiotopic and unaffected. Instead, it is the effective coordinates of the RFs that actually shift so that they are simultaneously retinotopic and logically spatiotopic. This leads to neural phenomena such as receptive field remapping (Duhamel, Colby, & Goldberg, 1992; Merriam, Genovese, & Colby, 2007). This shift compensation along with lateral visual flows will be responsible for the *temporal* cut & paste form of completion (Van Roosmalen, Kokaram, & Biemond, 1999).

The interaction between emergic units, i.e., of flows of information, is not governed by any equations. Each link simply delays information by a *tick* of time (nominally 10ms). Instead, it is the Emergic Unit that computes directly the functional interactions among flows, while the emergic port can sum visual values in a spatiotemporal manner. These functions are under a standard imperative programming paradigm. Therefore, no solutions to a set of governing equations are required leading to an emergic system that is robust and can be completely parameter free.

Functions within ECM and FACADE

Note that ECM's motion compensation functionality is not covered within FACADE which is geared to perceive object motion under static eyes. Similarly, FACADE has a function for the processing of apparent motion, which is not currently handled within ECM. So one cannot directly compare the overall complexity between FACADE and ECM at this level of analysis. Although it appears within Figure 1 that the

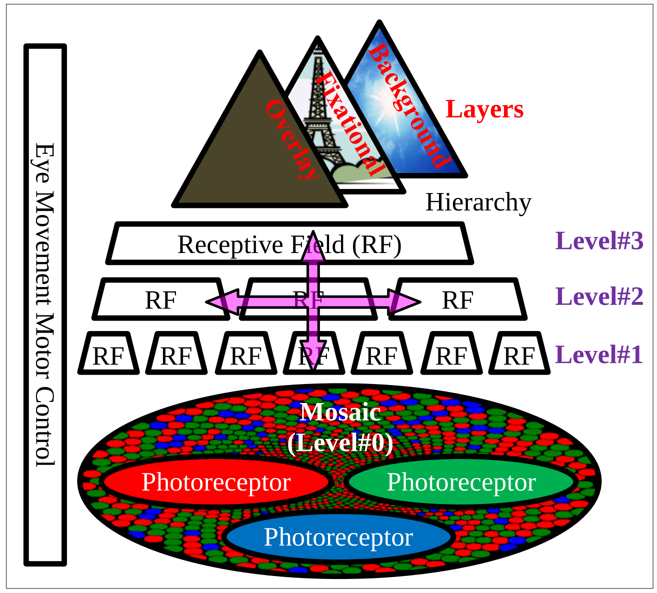


Figure 2: ECM Structural Summary

Emergic Units, which represent individual neurons in this case, are more complex than typical generic models of neurons, this is because the model incorporates dendritic processing along with spiking (see Leibovitz & West, 2013).

Each RF (represented by an Emergent Unit) has a single, trivial *handling missing data* function that is defined as follows:

1) If upward visual information is missing, use lateral information.
2) If it is still missing, use downward information.
3) Send out the information up, down and across the RF hierarchy, if available.

The difference between FACADE and ECM is over the concept of function and its attribution. At first glance, FACADE is more neurally detailed, but ECM attributes some functions to dendritic processing, which FACADE does not. FACADE has large grained functions to implement and control interactions between its Border Contour System and Feature Contour System, but these are highly sensitive to parameters (due to recurrence in the system). Overall, the functions in FACADE are large but not emergent (that is they arise from dedicated neural circuitry within a module). In contrast, ECM does not have explicit border processing, but acts as if it does. The fine-grained functions of ECM are explicit while higher level behaviours emerge.

Model of Narrow Slit Viewing

Anorthoscopic perception is perception under *abnormal* viewing conditions. One such scenario is when a wider view of the world is *perceived* than can be *sensed* at any one time when looking through a narrow slit. *Narrow slit viewing* and *aperture viewing* are common nicknames for this phenomenon. Somehow, visual information must be integrated across viewpoints.

An initial ECM model of anorthoscopic perception demonstrated positive results (Leibovitz, 2013b). A second test, using the stimulus viewed in Figure 3, provided the opportunity to analyze here, in greater detail, the spatial and temporal forms of border and surface completion, and to explain how these come about without corresponding modules to produce them.

Test Methodology

The ECM model was augmented to create a narrow slit viewing situation (Figure 4). The foveal blue scotoma as well as a central blind spot were also included in the model (Figure 5). The blue scotoma was aligned over the edges of a blue square to demonstrate border effects without border processing.

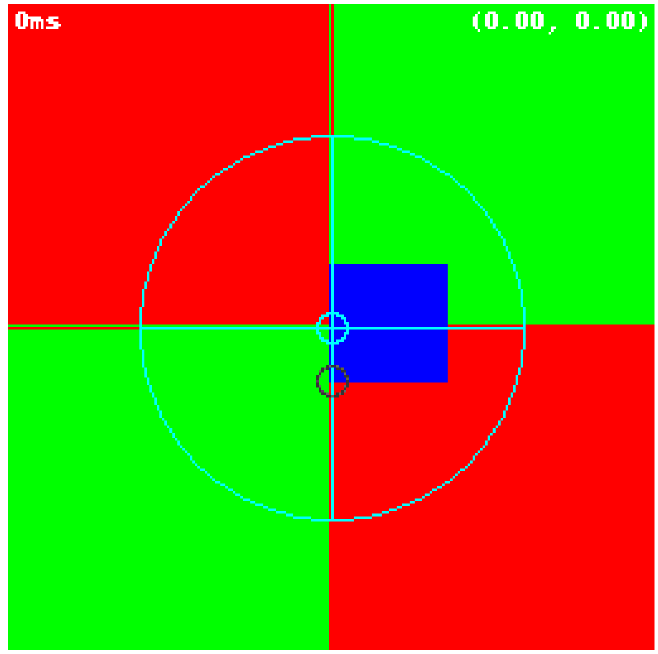


Figure 3: Test stimulus viewed @ 0ms

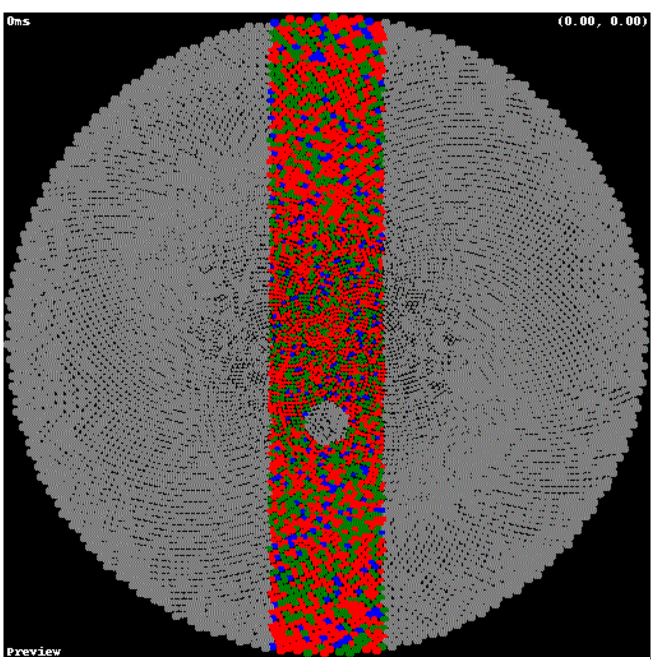


Figure 4: Agent's irregular retina (in False RGB) @ 0ms

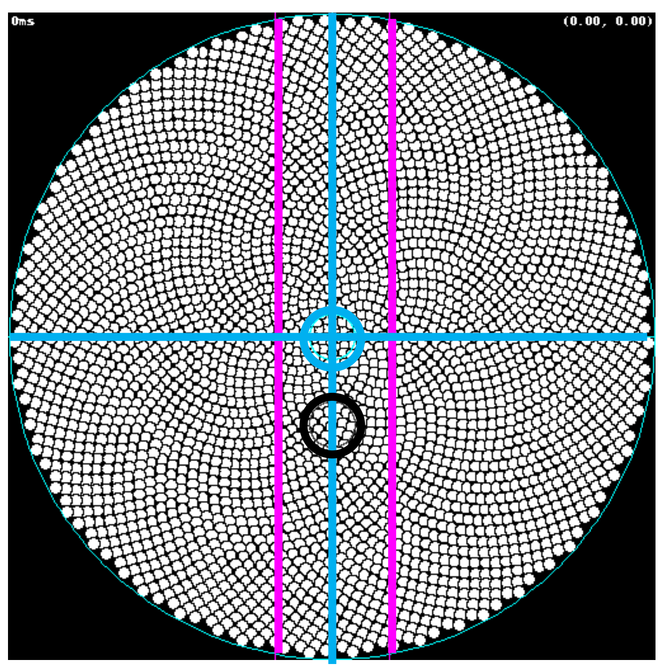


Figure 5: Agent's 1st level RF mosaic @0ms

Test Results

Extracts of the initial results relevant to this paper are shown within Table 1. Note that it takes 70ms before planned eye motion takes effect.

Table 1: Extract of Test Results

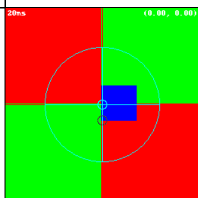
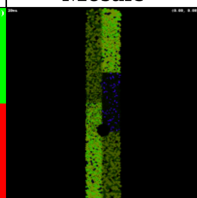
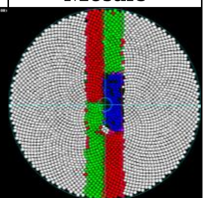
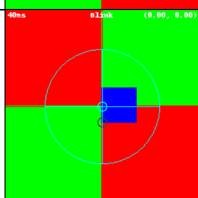
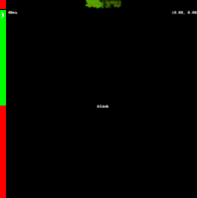
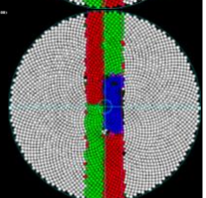
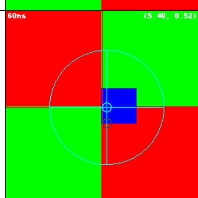
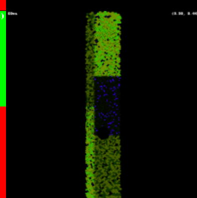
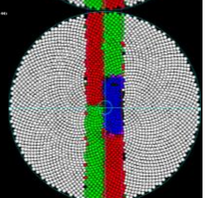
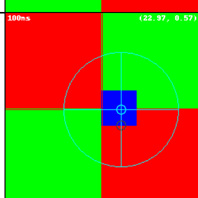
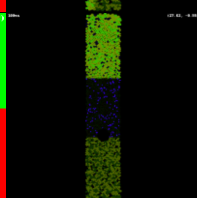
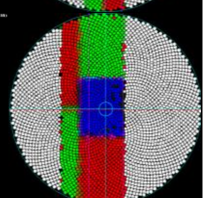
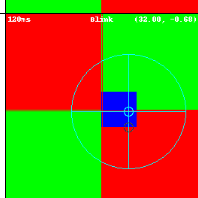
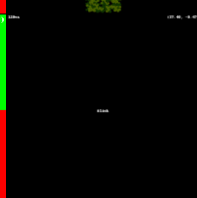
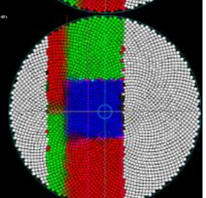
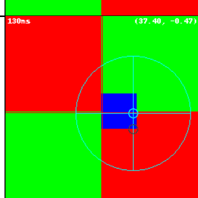
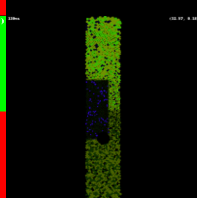
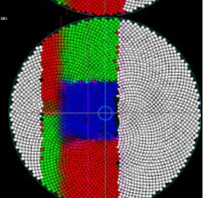
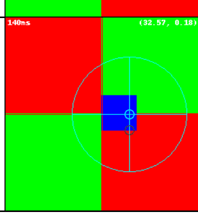
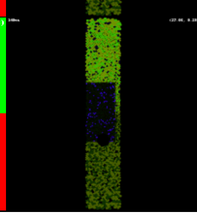
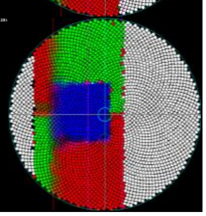
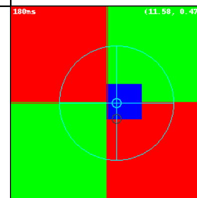
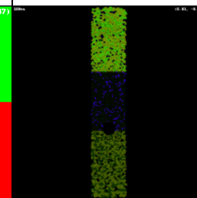
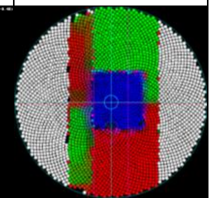
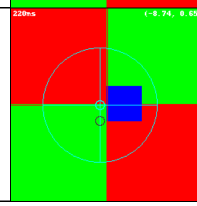
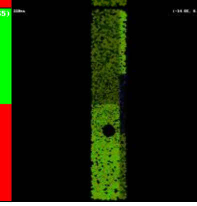
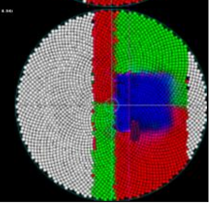
Time (ms)	Stimulus View	Photoreceptor Mosaic	1st Level RF Mosaic
20			
40			
60			
100			
120			
130			
140			
180			
220			

Table 1: Extract of Test Results

Detailed Analysis

Apparent Shifting of Flow

During times 70-150ms, the eye is moving rightwards and the flow of information within the RF mosaic is shifting leftwards to compensate. Recall that it is not the individual emergic values that are shifted, but the coordinates of each RF. After *handling missing data*, each RF broadcasts its output upwards, downwards and laterally within a local fan-out domain, much like in a local area network (LAN). Then each of the RF's own coordinates is actually shifted, so that the dendritic process within the incoming ports can sum up within the next tick only the newly appropriate visual values falling within the new coordinates of the RF.

During times 150-220ms, the eye is moving leftwards and the flow compensates by shifting rightwards. An apt metaphor for narrow slit viewing is that the retina or narrow view "paints" visual information onto a canvas. But rather than the retina moving over the canvas, the canvas moves itself. However, in the ECM model, there is no physical canvas made up of permanent spatiotopic neurons. Instead, the ephemeral flow of information constitutes the canvas. In essence, it is the "paint" itself that effectively moves, and it is the current neurons processing the paint that lends it such motion.

Spatial Completion Behaviour

It takes 60ms before planned eye movements take effect, so during that time, spatial completion respecting borders was demonstrated within the two scotomas. Figure 8 shows both scotomas initially as unfilled. The blind spot is missing all photoreceptor values and appears in white, while the foveal blue scotoma is just missing the blue photoreceptors and will appear as black over the blue portion of the square stimulus.

By 60ms, and before movement has started, both scotomas have been filled. Note that borders are respected. Neither red nor green take over portions within the blue scotoma. Only the blue seems to diffuse into the blue scotoma, but it does not extend into the regions where red and green are active.

Similarly, over the complete blind spot, filling-in seems to respect all borders.

Temporal Completion Behaviour

To test the temporal completion behavior a second simulation was run. This simulation was the same as the first except that the eyes blink, creating a temporal disruption in the flow of information. In addition, this simulation also included eye jitter, which is also a type of temporal disruption. In Figure 6, the eye is currently moving rightwards so information is being shifted leftwards. A blink occurs at 120ms that impacts the 1st Level RF mosaic one tick later. Thus at 130ms the entire RF represents a flowcentric memory in motion. It is still shifted, but without new data coming up from the photoreceptors, there is a set of RFs with missing (white) values within the narrow photoreceptor view as indicated. At 140ms, information is shifted directly into the two scotomas respecting borders. Indeed the information is two ticks old because of the blink. During all of this, eye jitter is occurring without any blurring. FACADE and most other models of vision do not include jitter because it causes blurring. In ECM, not only is blurring absent, the jitter also functions to help filling in.

Mechanism

Spatial Completion Mechanism

Spatial completion is best exemplified while the eye is not moving. It takes a finite amount of time due to the apparent filling-in from averaged top down flows – as information flows up it is spatiotemporally averaged. The size of the regions to be completed, as well as the size of the RFs over the affected region are the biggest factors affecting completion times (Leibovitz, 2013a).

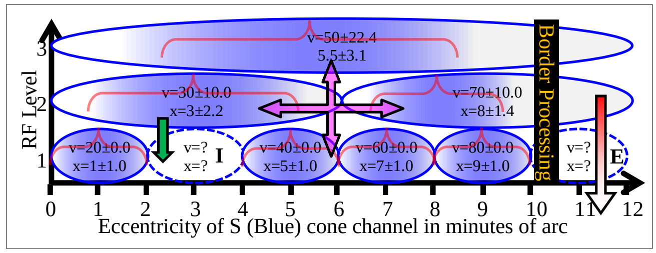


Figure 7: Interpolation vs. Extrapolation. Showing the mean of each RF value ± 1 standard deviation.

Emergic Border Processing

Unlike FACADE which has both static and motion border systems, as well as the contour portion of both the Border Contour System and the Feature Contour System, ECM has no explicit border or contour processing units nor assemblies. However, border processing emerges due to tagging values with their spatial extent. It is best summarized as the ability of RFs to interpolate but not to extrapolate as shown in Figure 7.

Basically, as visual information flows up the hierarchy, it is spatiotemporally averaged. For the blue (S) cones in the example, the amount of blue, and its standard deviation is automatically calculated. However, the same thing happens with the tags delimiting the spatial extent of each visual attribute. The figure shows the extent in only the horizontal dimension, and the calculated standard deviations serve as spatial delimiters.

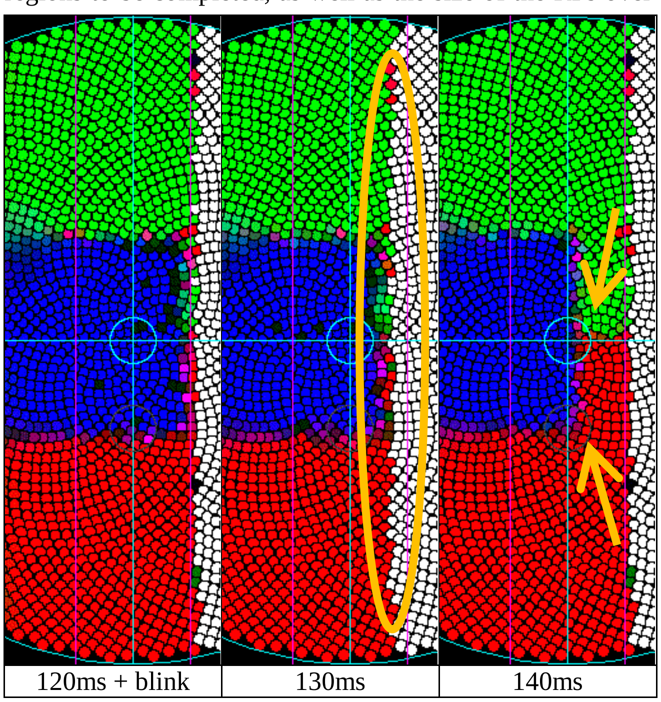


Figure 6: Temporal shifting into scotomas

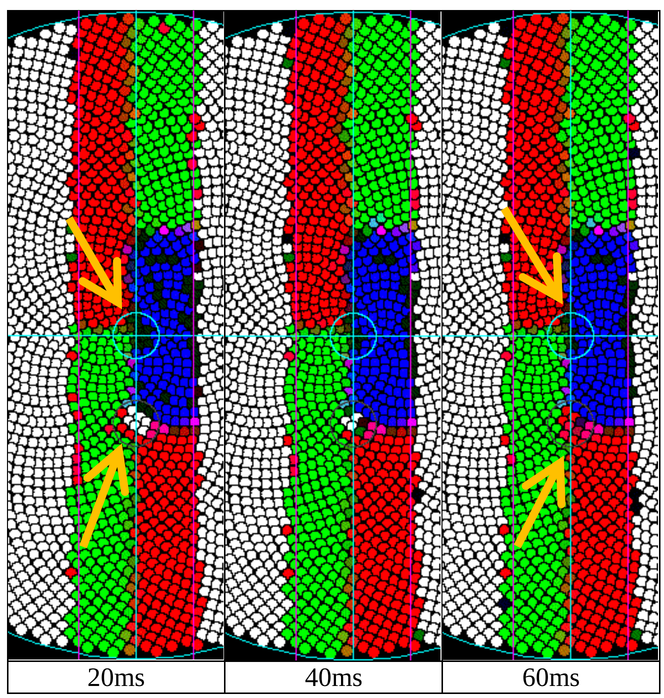


Figure 8: Spatial completion over scotomas

Thus the RF to be interpolated ('I') will accept the downward flow as it overlaps with its own spatial extent, while the RF to be extrapolated ('E') will reject the downward flow, effectively producing a border effect. There is neither diffusive filling-in nor border processing in an explicit manner.

This exemplifies a possible functional tradeoff between explicit high-level mechanisms and the lower-level representation of visual information with its spatiotopic coordinates. This low-level representation is associated with a physical tradeoff towards finer neurobiological structures.

Temporal Completion Mechanism

Temporal completion occurs instantaneously due to advanced shifting of information laterally. If the information coming up is somehow missing, possibly due to a blink or blind spot, then it can be retrieved immediately from lateral ports. Effectively, temporal completion is a *cut & paste* from the previous frame (Van Roosmalen et al., 1999). Information can even originate from earlier frames with longer intervening blinks or if shifting over larger scotomas.

Shifting or even temporal *cut & paste* do not involve diffusive filling-in of surfaces, nor border mechanisms. Because it is pre-emptive and instantaneous, edges always line up, even through scotomas. This prevents smear and differentiates flowcentric systems such as ECM from systems such as FACADE that are subject to smear under continuous motion. Essentially, systems such as FACADE treat neurons as pixels. Using the painting metaphor from above, these models assume a canvas of dedicated, fixed neurons, on which the visual system paints. These neurocentric systems cannot prevent smear under continuous motion as they retain visual states within neurons that do not compensate for motion (Leibovitz, 2013a). In contrast, the ECM model is flowcentric and does not retain a static representation of the visual field. Instead, visual information flows between neurons. Returning to the painting metaphor, it is as though the paint moves around on the canvas.

Summary

In FACADE and similar systems, high level functions are realized as separate modules constructed with relatively simple models of the individual neurons involved. These systems often require numerous parameters that can lead to brittleness. We have demonstrated that using more complex models of neurons can produce a more robust system, in which the phenomenological functionality that we observe is derived from the emergent properties of the dynamic interactions of lower level systems.

References

Berzhanskaya, J., Grossberg, S., & Mingolla, E. (2007). Laminar cortical dynamics of visual form and motion interactions during coherent object motion perception. *Spatial vision*, *20*(4), 337–95.

Duhamel, J.-R., Colby, C. L., & Goldberg, M. E. (1992). The updating of the representation of visual space in parietal cortex by intended eye movements. *Science*, *255*(5040), 90–2.

Grossberg, S. (2010). Neural Models of Motion Integration, Segmentation, and Probabilistic Decision-Making. In U. J. Ilg & G. S. Masson (Eds.), *Dynamics of Visual Motion Processing* (pp. 283–311). Boston, MA: Springer US.

Grossberg, S., Léveillé, J., & Versace, M. (2011). How do object reference frames and motion vector decomposition emerge in laminar cortical circuits? *Attention, perception & psychophysics*, *73*(4), 1147–70.

Grossberg, S., & Rudd, M. E. (1992). Cortical dynamics of visual motion perception: Short-range and long-range apparent motion. *Psychological Review*, *99*(1), 78–121.

Hu, D., Zhou, Z., & Wang, Z. (2011). Processing real-world imagery with FACADE-based approaches. *Frontiers of Electrical and Electronic Engineering in China*, *6*(1), 120–136.

Leibovitz, D. P. (2013a). *A Unified Cognitive Model of Visual Filling-In Based on an Emergic Network Architecture*. Carleton University.

Leibovitz, D. P. (2013b). Flowcentric Model of Narrow Slit Viewing in an Irregular Retina with Blind Spots, Eye Motion, Blinks and no Smear: Base Model Changes, Test Details and Test Results. *Dr. David Pierre Leibovitz*. Retrieved from http://dpleibovitz.upwize.com/?p=373

Leibovitz, D. P., & West, R. L. (2013). Dendritic+ Processing in an Emergic Network Model of Narrow Slit Viewing. *Proceedings of the 12th International Conference on Cognitive Modeling (ICCM 2013)*.

Merriam, E. P., Genovese, C. R., & Colby, C. L. (2007). Remapping in human visual cortex. *Journal of neurophysiology*, *97*(2), 1738–55.

Van Roosmalen, P. M. B., Kokaram, A. C., & Biemond, J. (1999). Fast high quality interpolation of missing data in image sequences using a controlled pasting scheme. In *Proceedings of the 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP99)* (pp. 3105–3108 vol.6). IEEE.