

A Review of Predictions From the Dynamic Routing Circuit

Kevin Leung

Stanford University

A Review of Predictions From the Dynamic Routing Circuit

In 1993, Olshausen, Anderson, and Van Essen presented the Dynamic Routing Circuit (DRC) as a model for how humans propagate an image of any size from anywhere on the retina into a single location for recognition in IT. The DRC proposes that the receptive fields of neurons in the visual pathway can shift and expand and shrink. Control neurons change the strength of connections between layers and focus attention through receptive field changes. In this paper, I will review the predictions from the DRC based on recent studies about receptive fields of neurons in visual areas.

Summary of the Dynamic Routing Circuit

The DRC is composed of 5 layers of nodes meant to represent V1a, V1b, V2, V4, and IT with each node corresponding to a population of neurons. To replicate any object in the visual field in IT, control neurons dynamically modulate the connection strengths between layers so a node only receives input from a subset of the nodes that feed into it. The combination of nodes activated forms a window of attention to encompass objects of different sizes at different locations.

To find all of the objects in an image, the DRC performs a 5-step process. First, it runs a low-pass filter over the image to generate a saliency map. Second, it selects a window of attention from the saliency map and maintains corresponding connections and weakens all others. Third, the selected object is compared with objects in memory in IT, and the exact window of attention is adjusted to perfectly match the memory. Fourth, the object is recognized and marked, and finally, the region is inhibited and the process begins again.

As mentioned, the network structure represents the visual areas in the “what” pathway. Each node has a receptive field that it responds to, and the modulation of connections represents

shifts, expansions, and contractions of that receptive field. Control units that control changes are thought to originate in the pulvinar. Roughly, the authors speculate that the saliency map exists in the posterior parietal cortex (PP), and connections from that to the superior colliculus to the pulvinar tell the control units where to focus attention.

All of these associations, however, are only predictions, and all are worth reviewing. Beyond the large structural predictions, another is that the receptive fields of neurons in the visual pathway should change depending on where attention is directed. Prior to the DRC, Moran and Desimone discovered that when 2 stimuli are presented in the classic receptive field (CRF) of a single neuron in V4, the response from the unattended stimuli is significantly attenuated, averaging firing rates of 36% of the baseline (1985, p. 783). Although this study suggests some change in the neuron's receptive field to focus attention, more specific measurements of responses at various locations test the scale of possible changes. Additionally, the DRC predicts that receptive fields should smoothly track along with movement of visual attention, and the size of receptive fields should also track the size of the window of attention.

Although these specific predictions haven't been tested, recent studies have considered the effects of changes in the receptive field and given some support to these ideas.

Effect of spatial attention on local area

In 1997, Connor et al. wanted to determine how the response of V4 neurons was affected by spatial and object-based attention. In their experiment, a Macaque monkey fixated at one point and attended a ring positioned near the receptive field of the neuron being monitored. Within the receptive field, they flashed oriented bars at different locations and found that the response changed significantly depending on the distance from the attended ring to the bar stimulus. The further the stimulus was from attention, the smaller the response. By comparing

responses when the monkey attended to opposite sides of the receptive field, they found an average fractional response shift of .33 in neurons with significant changes (p. 3205).

This result is not surprising based on predictions by the DRC. In the model, control units are represented by Gaussian distributions located over a set of connections. By establishing the window of attention using these control units as bases, the Gaussian distribution creates a gradient to the response of various populations near the window of attention. More importantly, they found a measurable shift in the receptive field of a neuron towards an area of attention. Although the results don't demonstrate any changes in size, they do support shifts more concretely than the earlier work by Moran and Desimone (1985).

For further work, I would be interested in a more fine-grained study where the relative distance of the stimuli and attended ring is changed by moving the ring. Another study discussed later will sample more locations in the receptive field with stimuli, but even that doesn't necessarily demonstrate smooth tracking of the receptive field with attention. This study would introduce additional confounding variables, particularly from neuronal tuning to motion, but one could hopefully find a neuron that is only tuned to the relevant features.

A final caveat of the study is that its presentation doesn't actually match the conditions of the model. In the DRC, receptive fields only change within the bounds of that population's classic receptive field. Because attention is focused outside of the classic receptive field of the neuron recorded, the study actually addresses a potentially harder, but not identical, problem. Even so, it does support receptive field shifts overall and explain what changes might occur outside of the window of attention.

Importance of features in tuning

In 2008, David et al. looked at the effects of features in the tuning of V4 neurons. We know that neurons not only respond within a receptive field but are also tuned to particular types of stimuli depending on its orientation or movement. This particular study looks at whether neurons exhibit spectral tuning in addition to spatial tuning, becoming more or less sensitive depending on the specific features of the target. With natural images, they found that spectral tuning changed significantly with feature-based attention but not with spatial attention.

This study supports the theory that matching occurs from the influence of top-down attention. Not only is IT looking for features, it's also recruiting other areas to focus as well. This effect supports the existence of feedback to modulate the response profiles of neurons in the visual pathway. So far, the evidence for control neurons and high-level control of neuronal responses has been scarce. David et al. conclude that feedback connections into V4 causes the spectral shifts, though they don't speculate from where the connections may originate. They also suggest that because feature-based attention and spatial attention have different effects on spectral tuning, separate mechanisms exist for these two types of attention.

Although these results don't affect the process of finding a window of attention in step 3 of the DRC, they are important for step 4. These feedback connections might exist for IT to adjust the receptive fields of neurons. Consider one possible interpretation of these results. The DRC uses Hopfield associative memory to match images, though the authors don't stand by it as the correct mechanism (Olshausen, Anderson, & Van Essen, 1993, p. 4706). Instead, spectral tuning shifts may resolve the exact contents of a window of attention. Not only does the adjustment move the window of attention, it may also increase the resolution of the image by focusing on features. Unfortunately, the lack of spectral tuning changes with spatial attention makes this theory less likely. It's possible that these effects might still occur; spectral tuning was

only measured using orientation and spatial frequency, and these neurons could be selecting for more complex features.

Spectral tuning from feature-based attention must also happen globally because it isn't spatial attention. Therefore, the results also provide an alternative explanation for how windows of attention are picked. The DRC only uses a low-pass filter for general saliency to pick blobs for additional processing for attention. Spectral tuning suggests that something far more complex may be occurring as response is enhanced for relevant features.

Expansion of objects in V1

Most of the previous studies on receptive fields share 2 qualities. First, they primarily record from V4, a higher level visual area where we can assume that generalization of visual information occurs. Second, receptive field changes seen tend to be shrinkage around a stimulus. A 2006 study by Murray, Boyaci, and Kersten addresses a very different question by looking at how objects are represented larger than reality in V1.

In their study, subjects saw a 2D image of 2 spheres of equal size drawn to appear as a 3D scene. Lines on the image provided depth cues so that the higher sphere appeared to be further away. Although both spheres covered an equal visual angle, subjects consistently reported that the back sphere appeared larger. In fMRI scans, the authors found that the eccentricity of the back sphere extended beyond that of the front sphere. Comparing responses in the regions of interest in V1, they found that the difference in response was similar to the difference between disks covering 6.5° and 8.125° (Murray, Boyaci, & Kersten, 2006, p. 430).

The authors concluded that the retinotopic map in V1 doesn't strictly maintain the size of images in the same part of the visual field. Given 3D cues, V1's map instead reflects the apparent size of an object in its environment. Although 3D information is ignored in the DRC,

this study does support the existence of more complicated mechanisms in the maps of visual areas.

The authors propose that higher visual areas provide feedback to V1 since depth cues require information over a much larger area than V1 can recognize. Such connections would do even more than what the DRC suggests: not only would they modulate the strength of feedforward connections to change the map of subsequent visual areas, they actually remap the location of objects in V1. If this proposed mechanism existed, the much smaller tweaking of matching a window of attention to a memory in IT becomes much easier.

As a more general mechanism, this study also shows how receptive fields might appear to grow for larger objects. Unlike the shrinking effects we've seen, the stimulus is represented larger than it actually is. Higher visual areas could maintain a level of invariance with respect to the size of an object in V1 based on these transformations. Consider this possible mechanism for how these changes might occur. An image, like the sphere in this study, is received and mapped in V1. Initially, it represents both spheres as the same size, but another area, such as IT, recognizes the cues for depth. One possible way to propagate this information back to V1 would be to simply tell V1 to remap its input, subsequently forcing every intermediary visual area to remap as well. Another more viable explanation might be that IT uses feedback connections in each intermediary area and adjusts receptive fields as it goes. In this manner, V2 and V4 can simply recognize the new map from V1 just by changing connection strengths instead of remapping.

This study, however, can also be interpreted as strong evidence against the DRC. Whereas the DRC claims that receptive fields account for scaling, the remapping supported by this study might enlarge objects in the map to achieve similar saliency. And as a final caveat, it's

notable that this study analyzes functionality at a much higher level than proposed by the DRC. Most of the other changes noted have been on single-cell recordings or small populations of neurons, while these studies were done via fMRI. Even so, these large features may provide insights into the result of this work.

Receptive field changes in MT

I mentioned above my concern about better resolution of responses after receptive field shifts and shrinkages, and a study by Womelsdorf, Anton-Erxleben, Pieper, and Treue directly tested the receptive field of MT neurons for the size of changes (2006). After measuring the size of the classic receptive field, they presented up to 52 other stimuli across the receptive field and discovered a significant, average shift of 3.0° along with a marginally significant shrinkage of 4.3% (pp. 1156-1157). A later study by the same group used a different model for firing rates that included attention as a multiplicative factor to the Gaussian model for the firing rates and found a more significant shrinkage of 11.2% (Womelsdorf, Anton-Erxleben, & Treue, 2008, p. 8938).

Although these results aren't surprising, they are still important. The small amount of shrinkage is somewhat alarming, especially in a high-level area like MT. Although the DRC doesn't make explicit predictions on the bounds of change in receptive field size, we would expect large changes to be possible because we can identify objects of many sizes, from the letter "a" on this page to the same "a" on a billboard. This study doesn't test anything so extreme, but 11.2% still seems small. It's possible, however, that populations of neurons may aggregate size changes for bigger results.

It's also important that these effects are in MT, which is not part of the "what" pathway where receptive field changes are supposed to happen in the DRC. In the DRC, the dorsal

pathway only creates the saliency map, though the authors don't speculate as to how that happens. Although the model doesn't strictly prohibit receptive field changes in the dorsal pathway, it also shouldn't be necessary. The exact same study in V4 instead of MT would be useful to not only directly support shrinkage in the ventral pathway, but also have specific data on the scale of changes. More generally, these similar shifts in receptive fields suggest that the two pathways may have significantly more similar functionality than previously thought.

Another promising note from their later work is that their model with multiplicative attentional gain is that the amplitude of neuron firing in a shrunk area actually increases to maintain an equal overall gain (Womelsdorf, Anton-Erxleben, & Treue, 2008, p. 8938). This increased firing might be the increased resolution predicted by shrinking the receptive field. These changes would be best measured by specifically recording neural responses to spatial frequency under modified receptive fields, though I am unaware of any work on this topic.

Discussion

Since the DRC was created, many studies have looked at the effects of attention on receptive fields in visual areas, and so far, the evidence seems to support that some dynamic control exists over the size and position of receptive fields. The shifts and size changes found have been small, especially compared to the possibility in the model of shrinkage up to 100 times smaller than the classic receptive field (Olshausen, Anderson, & Van Essen, 1993, p. 4712). Although the exact numbers aren't quite in-line, dynamic routing may happen on a much smaller scale than the model suggests, making it a significant but not core mechanism for visual attention.

Even though we have evidence for the predicted effects of the DRC, less evidence exists for the control neurons that drive the functionality of the model, and that may perhaps be the

most important aspect to explore. In designing the DRC, the authors had the Moran and Desimone study to inspire the receptive field shifts, and there has not been a corresponding paper for control neurons. One possible way to test this theory might be to do simultaneous recordings in both V4 and the pulvinar to find a correlation between receptive field shifts and activity from potential control neurons. If such a correlation were found, causation might be found through direct stimulation of the pulvinar and seeing what effects that has on attention. These studies, however, are largely speculative.

Even so, dynamic routing continues to be an important paradigm for understanding how these changes occur. The DRC takes a much higher level perspective on the visual system than many other models today: instead of focusing on Gabor filters and response profiles of individual neurons, populations of neurons are grouped together and assumed to faithfully reproduce the image. 12 years later, Olshausen maintains that dynamic routing is a key idea in driving theories about how V1 and the visual areas as a whole function (Olshausen & Field, 2005, pp. 1690-1691). So far, the evidence continues to support dynamic routing with more predictions left to be studied.

Bibliography

- Connor, C. E., Preddie, D. C., Gallant, J. L., & Van Essen, D. C. (1997). Spatial Attention Effects in Macaque Area V4. *The Journal of Neuroscience* , 17 (9), 3201-3214.
- David, S. V., Hayden, B. Y., Mazer, J. A., & Gallant, J. L. (2008). Attention to Stimulus Features Shifts Spectral Tuning of V4 Neurons during Natural Vision. *Neuron* , 59, 509-521.
- Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science* , 229 (4715), 782-784.
- Murray, S. O., Boyaci, H., & Kersten, D. (2006). The representation of perceived angular size in human primary visual cortex. *Nature Neuroscience* , 9 (3), 429-434.
- Olshausen, B. A., & Field, D. J. (2005). How close are we to understanding V1? *Neural Computation* , 17, 1665-1699.
- Olshausen, B. A., Anderson, C. H., & Van Essen, D. C. (1993). A Neurobiological Model of Visual Attention and Invariant Pattern Recognition Based on Dynamic Routing of Information. *The Journal of Neuroscience* , 13 (11), 4700-4719.
- Womelsdorf, T., Anton-Erxleben, K., & Treue, S. (2008). Receptive Field Shift and Shrinkage in Macaque Middle Temporal Area through Attentional Gain Modulation. *The Journal of Neuroscience* , 28 (36), 8934-8944.
- Womelsdorf, T., Anton-Erxleben, K., Pieper, F., & Treue, S. (2006). Dynamic shifts of visual receptive fields in cortical area MT by spatial attention. *Nature Neuroscience* , 9 (9), 1156-1160.