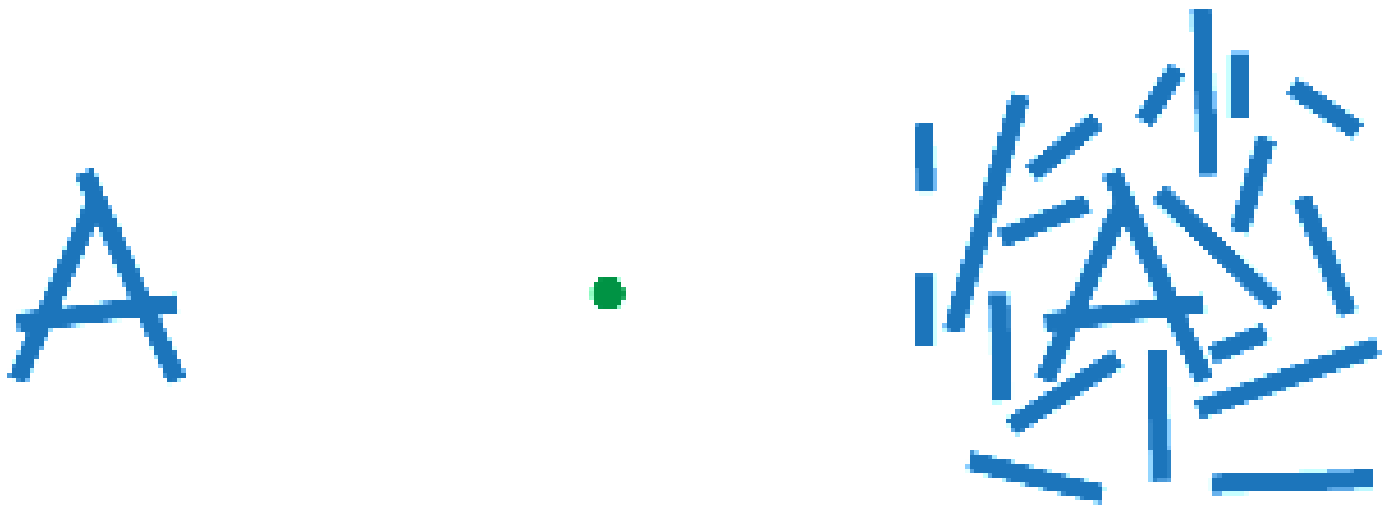


Modelling the limits of peripheral vision

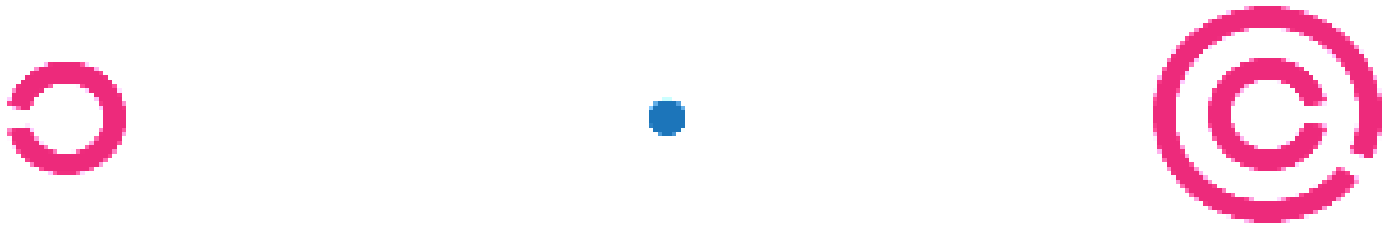
January 6, 2016

My latest publication about the limits of peripheral vision is now fully referenced and available. The journal has given me permission to host a PDF version for folks to download for a period of time, [so get it while it's hot!](#) (Be sure also to download the Supplemental Info!) The aim of this blog post is to describe the main points of the paper. **Preamble** Although we experience a highly detailed visual world, only a small portion of our vision – central vision – is high resolution. In the visual periphery, objects appear a bit blurry, and worse, they are prone to **visual crowding**. Visual crowding refers to the inability to recognise an object when other objects surround it. I've written about crowding before, [here](#) and [here](#), but for simplicity, check out the picture below. When fixating the central spot, the letter A on the left should be easily visible, whereas the same letter on the right is almost impossible to identify. (This demo was inspired by a similar one in a great review by Pelli and Tillman¹.)

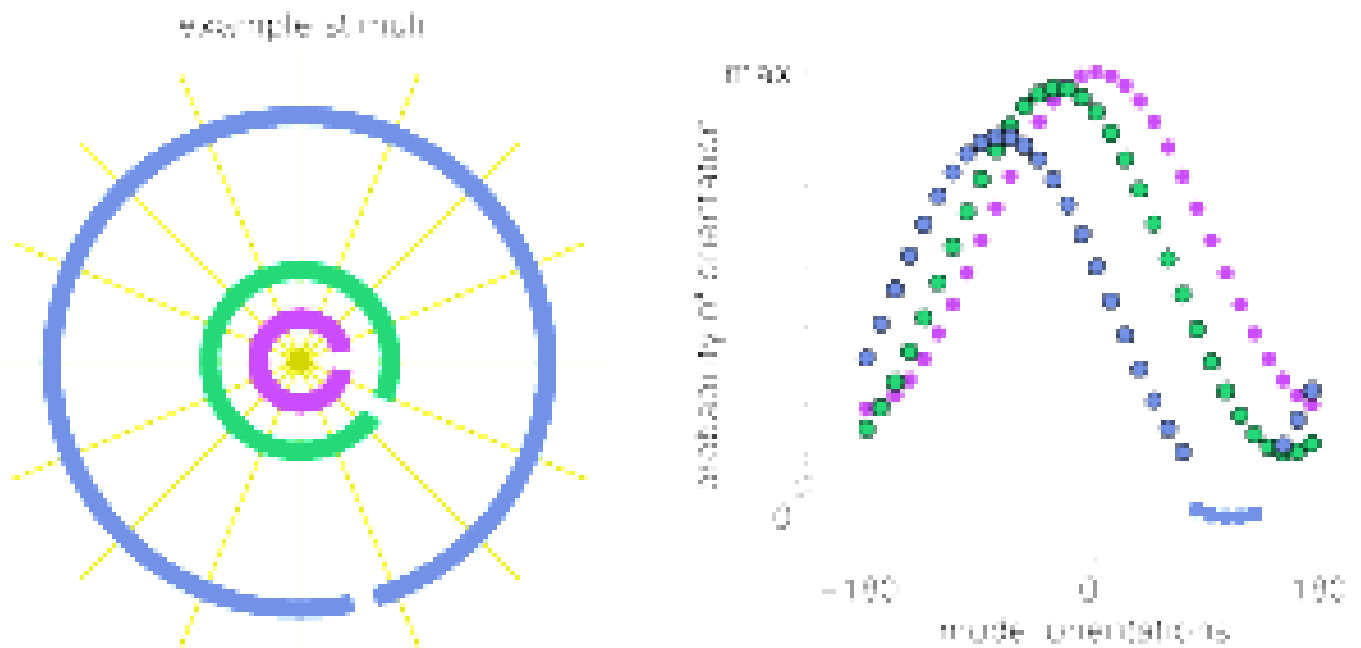


The obvious difference between the two sides of the above figure is that there are extra distracting lines on the right. These lines interfere with your ability to recognise the A. This is crowding. It happens with all kinds of objects, and happens throughout the entire visual field. *At any given moment, there are probably many objects in your vision that are crowded beyond recognition.* Right now, the speed at which you read this text is determined by your level of crowding². We tend not to be aware of such limited recognition, however, because [we can effortlessly make eye movements](#) to use our high-resolution central vision to break crowding. **Our contribution** In my paper, we discuss *why* crowding occurs. We describe a new method for measuring crowding, and we provide a computational model that simulates specific neural processes that may cause crowding. Our new method involves using a certain method of reporting the appearance of a stimulus in peripheral vision. Studies of crowding typically involve showing an observer a target, such as a letter, and then asking the observer to report the target identity. In these experiments, the observer's response can only be correct or incorrect. However, we know even from introspection when viewing demonstrations like the figure above, that when an object is crowded, we can still see *something*. In fact, being able to see something but not being able to correctly identify it is a defining feature of crowding³. It may be informative, therefore, to quantify crowded perception in more ways than as simply correct or incorrect. *We thus used a measure that allowed us to describe crowded perception as a loss of perception along a continuum.* **The task** We used a target stimulus like the one on the left below, known as a Landolt C. An observer saw this target in peripheral vision. Importantly, from trial to trial, it was rotated randomly, so that the gap section was orientated toward any direction. We then presented the observer with a second

Landolt C, this time in their central vision, that they could rotate by pressing buttons. They had to rotate the central C so that it matched the target orientation as closely as possible. On each trial, therefore, the observer's response was not classed as wrong or right, but instead we found their perceptual error – the difference in rotation between the actual target and the observer's report.



This method can be particularly useful for modelling, because over dozens of trials, the observer's errors conform very nicely to known mathematical functions. People with typical vision are generally pretty good at this – their errors cluster nicely around the actual target orientation. A recent study has related these sorts of perceptual errors to the noisy encoding of the stimulus in primary visual cortex⁴, the first cortical sight of visual processing (though in that study stimuli were presented at fixation, not in the periphery). We can quantify crowding by seeing how these patterns of data – and their corresponding mathematical functions – are affected in the case of a crowded target, like that shown on the right of the figure above. In this example, the target is surrounded by a larger ring with a gap in it. When fixating on the blue spot, you may be able to discern that there is a gap in each ring, but the exact position of the gaps may appear “fuzzy”. You may even experience the impression that gaps are sort of joined in some strange way, and that their orientations are very difficult to identify precisely. At the very least, it's unlikely the target gap will appear as clearly as it does on the left of the figure, despite the physical properties of the targets being the same. **Our human data** When we examined the errors that observers made, we noted that we could see in our data patterns that corresponded to separate, previously competing explanations of crowding. First, we found that, when the orientation of the target and distractor were similar, observers tended to report an orientation close to the average of the orientations. Such feature averaging has been reported widely⁵, and strongly supports the idea that the visual system simplifies perception in the periphery by finding higher-level statistical associations between objects⁶. Second, we found that, when the orientation of the target and distractor were dissimilar, observers reported either the target or distractor, and their reports in these cases were quite precise. These sorts of errors are known as confusions or substitutions⁷, because the observer can apparently see both gap elements of the stimulus, but they are apparently unsure about which gap belongs to which ring. You may be able to experience both of these during different periods of viewing the target stimuli above. Note that these effects disappear in our data when the distractor ring is large and very far away from the target, showing that these phenomena are linked to processes that depend on the spatial positions of items. (Note that there is nothing particularly new about these data – these patterns merely replicate the same effects that have already been widely reported across studies.) **The model** We next generated a computational model that simulates a basic visual process – the coding of orientation information. Because the human data in the uncrowded condition of our experiment conformed well to a mathematical function (a circular normal distribution), we simply asserted that there is some neural process that results in an output of data with the same probabilities as given by that function. That is, for any given uncrowded target, we make our model make the same sorts of errors humans make. We then make the model spatial, by asserting that it'll respond most strongly to objects centred right on the target; the strength of the model response decreases gradually with the distance of an object from the centre of the target. In the figure below, I've tried to give an intuition for how the model works. On the left are three example stimuli, the target is coloured magenta and the distractors are coloured green and blue. The yellow lines can be thought of as our model trying to detect whether there is a gap section at any of those orientations.



On the right side of the above figure are illustrative model responses to the stimuli presented on the left. For any given gap orientation the model will indicate the probability that various orientations of the gap are present. The colours of each of the distributions correspond to the stimuli on the left. The magenta target distribution is centred on 0° (the target orientation in [polar coordinates](#)), and is strongest because the target is closest to the centre of the model. You can see that the peaks of the distractor distributions are shifted away from zero, and are weaker than the target. **The model works...** In our study, we simply summed the model response to a target with the model's response to a distractor, and tested how well those summed probabilities predicted observers' perceptual reports. We essentially had our model perform the exact same trials our observers performed. This worked very well: the model predicted the average errors made by observers over a range of distractor conditions. More importantly, this model predicted the mixed pattern of data described above, in which some trials indicate that the observer reported the average of the target and distractor orientations, and in other trials they confuse which gap corresponds to which object. **The model works... sort of.** In a second experiment, we used a novel behavioural paradigm to quantify the level of detail an observer could identify. This experiment requires much more explanation, so I'll save it for another post. However, it's important to note that, although we found it preferable than models proposed by others, our model had to be altered slightly to be able to account for our observations in this second experiment. This issue – how well our model can be used to describe other datasets – is important and remains to be tested. So, although our model describes the data using our specific task, how well it represents what the brain is doing in crowded scenes requires much further testing. Based on another great model that is conceptually similar {vandenBerg:2010bj}, I'm optimistic these sorts of models can go far. **And finally...** As with all of science, our work builds upon a lot of great stuff of work that came before. While the content of the paper is new, similar methods⁸⁻¹⁰ and models^{11,12} have already been published. I'll save a discussion of the differences between our work and this other work in another post.

My paper's citation:

References

1. Pelli, D. G. & Tillman, K. A. The uncrowded window of object recognition. *Nature Neuroscience* **11**, 1129–1135 (2008).
2. Kwon, M., Legge, G. E. & Dubbels, B. R. Developmental changes in the visual span for reading. *Vision Research* **47**, 2889–2900 (2007).
3. Pelli, D. G., Palomares, M. & Majaj, N. J. Crowding is unlike ordinary masking: distinguishing feature

- integration from detection. *Journal of Vision* **4**, 1136–1169 (2004).
4. van Bergen, R. S., Ji Ma, W., Pratte, M. S. & Jehee, J. F. M. Sensory uncertainty decoded from visual cortex predicts behavior. *Nature Neuroscience* (2015). doi:10.1038/nn.4150
 5. Greenwood, J. A., Bex, P. J. & Dakin, S. C. Positional averaging explains crowding with letter-like stimuli. *Proceedings of the National Academy of Sciences* **106**, 13130–13135 (2009).
 6. Freeman, J. & Simoncelli, E. P. Metamers of the ventral stream. *Nature Neuroscience* **14**, 1195–1201 (2011).
 7. Strasburger, H. & Malania, M. Source confusion is a major cause of crowding. *Journal of Vision* **13**, (2013).
 8. Ester, E. F., Klee, D. & Awh, E. Visual crowding cannot be wholly explained by feature pooling. *Journal of Experimental Psychology: Human Perception and Performance* **40**, 1022–1033 (2014).
 9. Ester, E. F., Zilber, E. & Serences, J. T. Substitution and pooling in visual crowding induced by similar and dissimilar distractors. *Journal of Vision* **15**, 1–12 (2015).
 10. Tamber-Rosenau, B. J., Fintzi, A. R. & Marois, R. Crowding in Visual Working Memory Reveals Its Spatial Resolution and the Nature of Its Representations. *Psychological Science* (2015). doi:10.1177/0956797615592394
 11. van den Berg, R., Roerdink, J. B. T. M. & Cornelissen, F. W. A neurophysiologically plausible population code model for feature integration explains visual crowding. *PLoS Computational Biology* **6**, e1000646 (2010).
 12. Dakin, S. C., Cass, J., Greenwood, J. A. & Bex, P. J. Probabilistic, positional averaging predicts object-level crowding effects with letter-like stimuli. *Journal of Vision* **10**, 14 (2010).