



# Visual crowding illustrates the inadequacy of local vs. global and feedforward vs. feedback distinctions in modeling visual perception

Aaron M. Clarke<sup>1\*</sup>, Michael H. Herzog<sup>1</sup> and Gregory Francis<sup>1,2</sup>

<sup>1</sup> Laboratory of Psychophysics, Brain, Mind Institute, Science Vie, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

<sup>2</sup> Department of Psychological Sciences, Purdue University, West Lafayette, IN, USA

## Edited by:

Hulusi Kafaligonul, Bilkent University, Turkey

## Reviewed by:

Thomas S. A. Wallis, The University of Tübingen, Germany  
Ronald Van Den Berg, University of Cambridge, UK

## \*Correspondence:

Aaron M. Clarke, Laboratory of Psychophysics, Brain, Mind Institute, Science Vie, École Polytechnique Fédérale de Lausanne, Station 19, CH-1015 Lausanne, Switzerland  
e-mail: aaron.clarke@epfl.ch

Experimentalists tend to classify models of visual perception as being either local or global, and involving either feedforward or feedback processing. We argue that these distinctions are not as helpful as they might appear, and we illustrate these issues by analyzing models of visual crowding as an example. Recent studies have argued that crowding cannot be explained by purely local processing, but that instead, global factors such as perceptual grouping are crucial. Theories of perceptual grouping, in turn, often invoke feedback connections as a way to account for their global properties. We examined three types of crowding models that are representative of global processing models, and two of which employ feedback processing: a model based on Fourier filtering, a feedback neural network, and a specific feedback neural architecture that explicitly models perceptual grouping. Simulations demonstrate that crucial empirical findings are not accounted for by any of the models. We conclude that empirical investigations that reject a local or feedforward architecture offer almost no constraints for model construction, as there are an uncountable number of global and feedback systems. We propose that the identification of a system as being local or global and feedforward or feedback is less important than the identification of a system's computational details. Only the latter information can provide constraints on model development and promote quantitative explanations of complex phenomena.

**Keywords:** feed-forward, hierarchical models, feedback, object recognition, scene processing

## 1. INTRODUCTION

A common approach to understanding vision is to identify whether a particular aspect of visual perception involves “local” or “global” processing. Local processing suggests that the information needed for some behavioral task is determined predominantly by information that is spatially close to the target stimulus. Global processing suggests that information processing is influenced by elements that may be distant from the target. Distinguishing between visual processing as being local or global has long been an important aspect of the Gestalt approach to perception (see the review by Wagemans et al., 2012). The local vs. global distinction also plays an important role in characterizing the flow of information in visual cortex (e.g., Altmann et al., 2003) and identifying the order of processing for natural scenes (e.g., Rasche and Koch, 2002; Cesarei and Loftus, 2011).

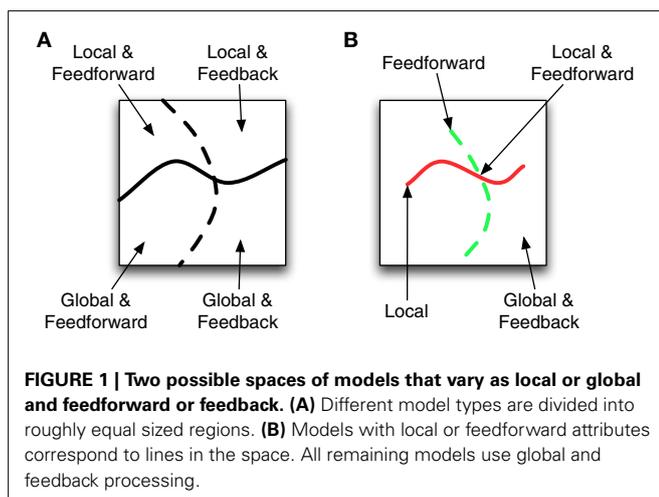
Likewise, many investigations try to identify whether visual processing involves “feedforward” or “feedback” processing. In a feedforward system the information flows in one direction, while in a feedback system the information flowing back and forth within and between areas can alter the processing at a given cortical location. In neuroanatomical studies, feedback processing is sometimes referred to as recurrent processing or re-entrant processing (especially when it involves information from higher

cortical areas projecting to lower visual areas). Since feedforward processing tends to be easier to model, interpret, and compute than feedback processing, it is often the starting point for computational and neurophysiological theories and serves as a standard comparison for subsequent studies that explore feedback effects. For example, Hubel and Wiesel (1962) proposed a local feedforward model that accounted for the properties of simple and complex cell receptive fields, and subsequent studies then proposed the existence of non-classical receptive fields by demonstrating effects of feedback or global processing (e.g., Von der Heydt et al., 1984; DeAngelis et al., 1994; Freeman et al., 2001; Harrison et al., 2007). Likewise, a popular theory of visual processing proposed that both a rapid feedforward sweep and a slower recurrent process is involved in different behavioral tasks to different degrees (Lamme and Roelfsema, 2000; Lamme, 2006), and many studies have explored whether particular phenomena depend on one or the other processing approach. Examples include Altmann et al. (2003) reporting evidence for feedback processing in an fMRI study of perceptual organization; Enns and Di Lollo (2000) arguing that some forms of visual masking require re-entrant signals that represent objects; Juan and Walsh (2003) using TMS to argue that the representation of information in area V1 is influenced by feedback from other areas; and Keil et al. (2009) reaching a

similar conclusion for emotionally arousing stimuli using an ERP study.

Experimental vision science is full of many other examples of investigations into local vs. global and feedforward vs. feedback processing, and we generally agree with their methods and conclusions. However, we are less convinced that these characterizations are especially useful for developing models of visual perception that might account for observed behavioral phenomena, and we suspect that the benefits of the local vs. global and feedforward vs. feedback dichotomies have been somewhat overstated. The seeming appeal of investigations that distinguish between local vs. global and feedforward vs. feedback processing may derive from a misunderstanding about the general properties of complex systems. **Figure 1A** schematizes one way of conceptualizing model space. The solid wavy line separates local models from global models while the dashed line separates feedforward models from feedback models. Under such a model space, identifying whether a system requires local or global processing divides the possible number of models nearly in half. Likewise identifying whether a system requires feedforward or feedback processing again divides the number of possible models in half. If the model space were as dichotomous as in **Figure 1A**, then investigations about the local vs. global or feedforward vs. feedback nature of visual processing would be very beneficial to modelers.

However, the characterization in **Figure 1A** cannot be correct because there must necessarily be fewer feedforward and local systems than feedback or global systems (e.g., every feedforward system can be augmented with multiple types of feedback), so the model space depicted in **Figure 1B** is closer to reality. Here the local models are characterized by a thin red line and the feedforward models are characterized by a thin dashed green line. The class of local and feedforward models is the small intersection of these lines, while global and feedback models correspond to almost everything else. If this perspective of the model space is correct, then scientists gain a lot of information by knowing a system uses local (Weisstein, 1968) or feedforward processing (VanRullen et al., 2001), but they gain very little information by knowing the model uses global and feedback processing.



Our argument is not that distinctions between local and global or feedforward and feedback processing provide *no* information about the properties of the visual system; but if **Figure 1B** is correct, then such distinctions will not generally provide sufficient constraints to promote model development for the identified effects. While this limitation may already be clear to many modelers, it seems that some experimentalists do not fully understand that such distinctions provide very little guidance for model development. Part of the problem is the underlying textbook assumption that there is *one* standard feedforward model and another standard feedback model, which implies that all we have to do is perform an experiment to see which type of model better describes task performance. It is indeed true that there are successful and popular feedforward and feedback models. The feedforward model of Riesenhuber and Poggio (1999), for example, has been used successfully for things like fast-feedforward object recognition or scene classification (e.g., Hung et al., 2005; Serre et al., 2005, 2007a,b; Poggio et al., 2013). Similarly, the feedback model of Grossberg (e.g., Grossberg and Mingolla, 1985) has spawned a multitude of subsequent publications (e.g., Grossberg and Todorovic, 1988; Grossberg and Rudd, 1989; Grossberg, 1990; Francis et al., 1994; Francis and Grossberg, 1995; Dresch and Grossberg, 1997; Grossberg, 2003; Grossberg and Howe, 2003; Grossberg and Yazdanbakhsh, 2003; Grossberg et al., 2011; Foley et al., 2012). Clearly there is an important role for both types of model architectures. However, the success of these models is not simply because of their feedforward or feedback architecture. Even these “popular” models involve parameter variations and additional stages from one paper to the next that make them suitable for modeling one experimental data set, but not another. Moreover, there exists a broad continuum of models that are designed to model various phenomena and include various amounts of feedforward and feedback processing, or local and global processing, and that are all different. In this sense, there is not really a “standard” model for the visual system. Even V1 receptive field models are vast and varied, including such models as Gabor (Gabor, 1946; Jones and Palmer, 1987), balanced Gabors (Cope et al., 2008, 2009), difference of Gaussians (Sceniak et al., 1999), oriented difference of Gaussians (Blakeslee and McCourt, 2004; Blakeslee et al., 2005), the log-Gaussian in the Fourier domain (Field, 1987), and many more, all of which produce similar, but distinctly different effects when applied to natural images and lab illusions. Moreover, V1 receptive fields comprise just the first step in a model of visual cortex. Thus, no “standard” models exists for either feedforward or for feedback architectures, and similarly for a local or a more global connection architecture. Simply specifying one or the other type of architecture is not helpful for many modeling projects. To demonstrate our point, we consider empirical data from studies of visual crowding that show a clear non-local effect, and that likely require feedback mechanisms to enable perceptual grouping. We then describe the properties of three plausible models: one that can be considered to be feedforward and global, one that can be considered to be feedback and global, and one that can be considered to be feedback and global with a clear interpretation of perceptual grouping. We show through computer simulations that none of these models can account for the empirical findings

that motivated them. This result suggests that we need to stop focusing on unhelpful dichotomies such as local vs. global and feedforward vs. feedback and instead should explore other properties of visual perception that help identify robust computational principles.

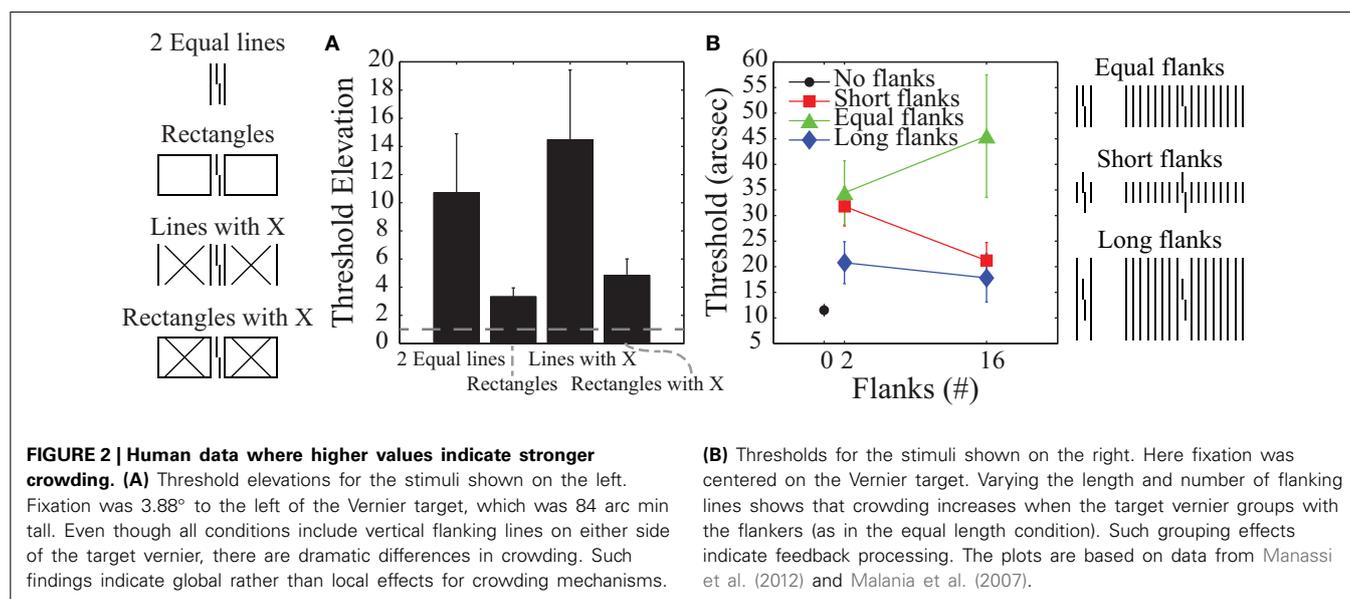
## 2. VISUAL CROWDING AS AN EXAMPLE

In visual crowding the discrimination of a target stimulus is impaired by the presence of neighboring elements. Crowding is ubiquitous in human environments. Even while you read these words, the letters appearing in the periphery of your visual field are crowded and largely unintelligible. Crowding can even be life-threatening in driving situations where a pedestrian can become unidentifiable by standing amongst other elements in the visual scene (Whitney and Levi, 2011). Moreover, visual crowding has been used to investigate many other aspects of perceptual and cognitive processing including visual acuity (Atkinson et al., 1988), neural competition (Keyser and Perrett, 2002), and awareness (Wallis and Bex, 2011).

The most popular models of crowding are local and feedforward models in which deteriorated target processing is due to information about the target being pooled with information about the flankers (e.g., Parkes et al., 2001). Although such pooling mechanisms are the default interpretation of crowding effects, recent studies have suggested that crowding involves global (rather than local) and feedback (rather than feedforward) processing (Maliania et al., 2007; Levi and Carney, 2009; Sayim et al., 2010; Livne and Sagi, 2011; Manassi et al., 2012, 2013). **Figure 2** schematizes eleven different types of stimuli where the task is always to identify the offset direction of a central target vernier. **Figure 2A** shows human vernier offset discrimination threshold elevations (relative to a no-flanker case), where larger threshold elevations indicate more crowding (from Manassi et al., 2012). The stimuli used are depicted on the far left-hand side of the figure. In all cases, the vernier is flanked by two vertical lines whose length matches the vertical extent of the vernier. The data

in **Figure 2A** indicate that the different flanker types do not produce equivalent crowding despite the identical neighboring lines. Although the flanking lines alone or with an “X” produce substantial crowding, there is very little crowding when the very same lines are part of a larger structure. A “local” mechanism, such as pooling, would predict similar (or stronger) crowding with the additional contours in the rectangle configurations. The observed decrease in crowding suggests that the phenomenon cannot be explained by local interactions between stimuli.

**Figure 2B** shows human vernier thresholds (Maliania et al., 2007) that have also been used to argue for feedback processing. Different experimental conditions varied the lengths of the flanking lines (shorter than, equal to, or longer than the vernier) and the number of flanking lines (0, 2, or 16). For the equal-length flankers, an increase in the number of flankers leads to stronger crowding, while for the short- and long-flanker lines, an increase in the number of flankers either reduced crowding or produced essentially no change. The argument for feedback processing has two parts. First, the data for the different conditions in **Figure 2B** suggest that crowding is strongest when the target vernier perceptually groups with the flankers (e.g., 16 equal-length flankers) and it is weakest when the target is perceptually segmented from the flankers (e.g., 16 short or long flankers). A sense of these grouping effects can be gained by looking at the schematized stimuli at the far right of **Figure 2B**. Second, perceptual grouping seems to require systems with feedback processing (e.g., Grossberg and Mingolla, 1985; Herzog et al., 2003; Craft et al., 2007; Hermens et al., 2008; Francis, 2009; Kogo et al., 2010). In particular, as Manassi et al. (2013) noted, the properties of crowding seem to defy low-level feedforward models based on stimulus energy or similar concepts (although they did not attempt to model their results). In their experiments they had subjects perform Vernier offset discrimination tasks and showed that when holding local information constant, global stimulus information still influenced thresholds. Thus, local information must have been propagated globally. Further experiments showed that



this local-to-global information propagation takes time, implying feedback and recurrent processing.

Since the crowding data in **Figure 2** indicate a role for global rather than local and feedback rather than feedforward processing, we wanted to use this knowledge to help develop a model of visual processing that accounted for crowding. Several models for crowding exist in the literature (e.g., Wilkinson et al., 1997; Balas et al., 2009; Greenwood et al., 2009; Van den Berg et al., 2010; Freeman and Simoncelli, 2011). Our intent here is not to classify these models as feedforward/feedback or local/global, and see how well they work, but rather to examine some clear examples of models using various amounts of feedforward/feedback and local/global processing and demonstrate the utility (or lack thereof) of knowing that a phenomena requires feedforward/feedback or local/global processing for modeling behavioral results. As the following sections demonstrate, we found this knowledge to be inadequate, and we believe that the modeling challenges here reflect issues that also apply to other phenomena and modeling efforts. Although it is possible that we happened to simulate models that are poor fits for the phenomena, we deliberately investigated models that have successfully modeled similar stimuli and phenomena, so we believed that they might also be able to account for the empirically observed crowding effects.

### 3. A FEEDFORWARD GLOBAL MODEL: FOURIER ANALYSIS

Researchers have suggested that it is useful to describe visual processing in terms of Fourier components (Campbell and Robson, 1968; De Valois et al., 1982). Luminance values at different  $(x, y)$  coordinates in the pixel plane can be converted to weights for different sine wave frequencies' amplitudes and phases. In principle, such a transformation does not lose any information, so if the luminance image contains information about the offset direction of a vernier, then so does the Fourier representation. However, such information can become degraded or lost when important frequencies or phases are filtered out of the representation. Such filtering can be justified on neurophysiological grounds (Campbell and Robson, 1968; De Valois et al., 1982) or be chosen to explain perceptual phenomena. For example, multi-scale filtering can explain a variety of brightness illusions (Blakeslee and McCourt, 1999, 2001, 2004; Blakeslee et al., 2005).

Fourier decomposition can be considered to be a feedforward process, with a bank of filters that are tuned to different frequencies, orientations, and phases (Fourier, 1822), and such an interpretation is a common first-approximation to cortical visual processing (Campbell and Robson, 1968). On the other hand, Fourier analysis is decidedly global rather than local in the sense that the weights assigned to different frequencies are based on the pattern of luminance values across the entire image plane (Rasche and Koch, 2002; Cesarei and Loftus, 2011). It is also global in the sense that a filter that suppresses some frequencies will influence representations of luminance values across the entire image plane when the frequency weights are converted back to an image representation.

We developed a model that applies a Fourier analysis to the image, filters out a subset of spatial frequencies, applies a Fourier synthesis to construct a filtered version of the image, and then compares the output with a template for discriminating right-

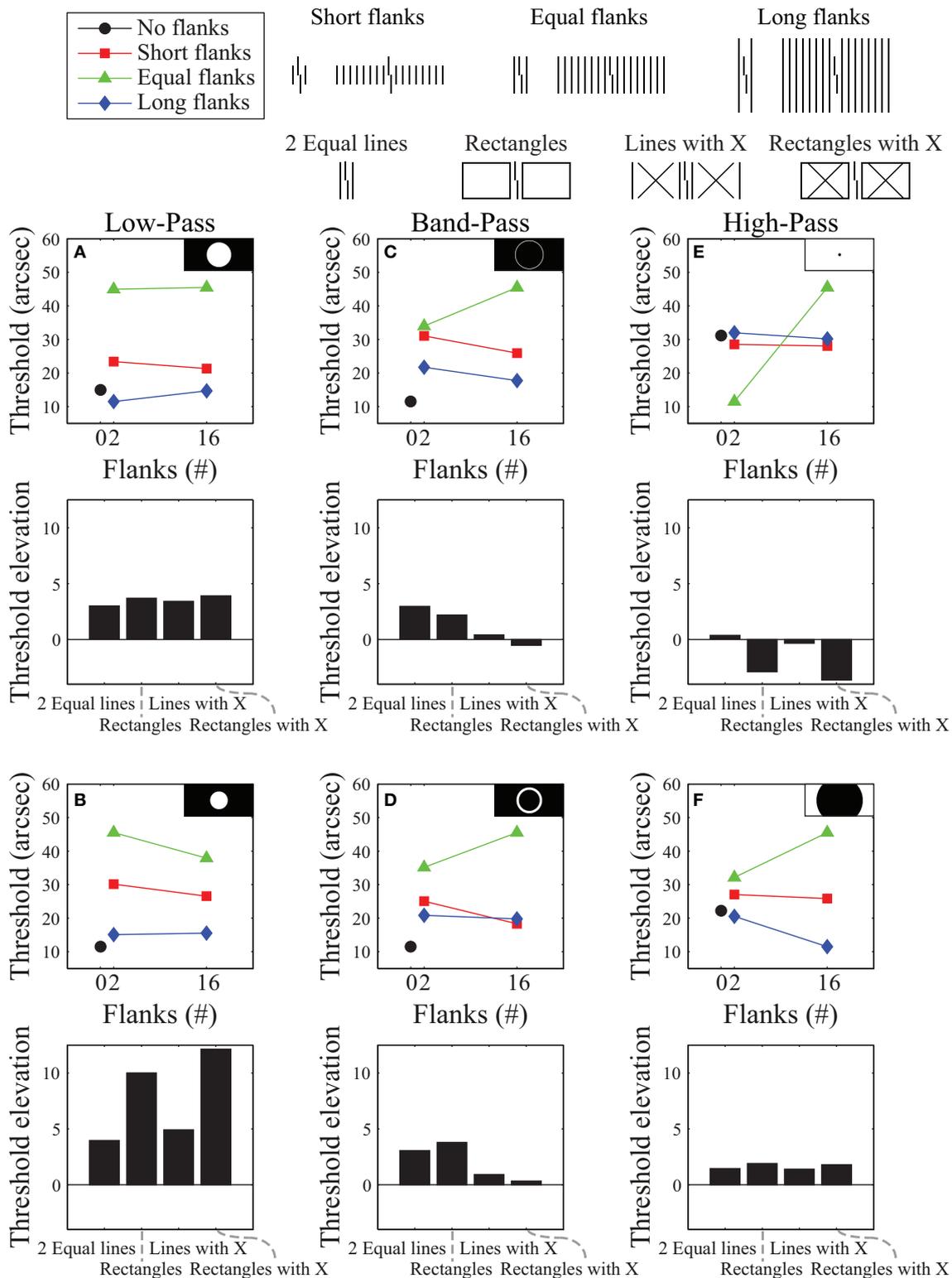
from left-offset verniers. The difference between template matching results for the left- and right-offset verniers are subtracted and the difference is then inverted and linearly scaled to the range of the human data. Model details are provided in the Supplementary Material. To try to match the empirical data, we examined various filtering schemes, including high-pass filtering, band-pass filtering, and low-pass filtering. **Figure 3** shows a representative selection of results for the stimuli used to produce the data in **Figure 2**. Even though they all allow for global processing, many of these frequency filtering functions produce results that differ dramatically from the human data. Within each filtering scheme we identified the filter parameters that yielded the smallest sum of squared residuals between the model and human data from **Figure 2** by exhaustive, brute-force search over the entire parameter space. **Figure 3C**, shows the best fit overall, which was obtained with a band-pass filter.

This best filter mask for the data used in **Figure 3C** does a reasonably good job at reproducing the human data from **Figure 2B**, but it does a poor job reproducing the human data shown in **Figure 2A**. Although the model roughly follows the pattern of the data for the two-line flanker and rectangle conditions, it predicts very little threshold elevation (and even threshold improvement) for the conditions with an "X" superimposed over the flanker regions. These predictions do not match the empirical data. The other filter functions also fail to reproduce the human data for these flanker conditions.

Moreover, the best filter is fragile in that small changes in bandwidth and/or center frequency lead to very different model predictions. This fragility is demonstrated in **Figure 4**, which shows model performance for band pass filters that are only slightly different from the filter that produces the best fit to the empirical data in **Figure 2B**. This behavior is surprising since Fourier models generally tend to fail gracefully with small deviations from the optimal filter parameters. The wildly varying model behavior suggests that the good fit exhibited in **Figure 3C** reflects over-fitting rather than a mechanistic explanation of the behavior. Overall, the model fails to account for the human data in a robust way. Such a failure occurs even though the model is inherently global in terms of processing, and thus satisfies one of the requirements seemingly needed to account for crowding effects. We cannot definitively claim that all Fourier-type models cannot account for crowding effects, but it seems that a good model does not easily appear simply because it has global processing.

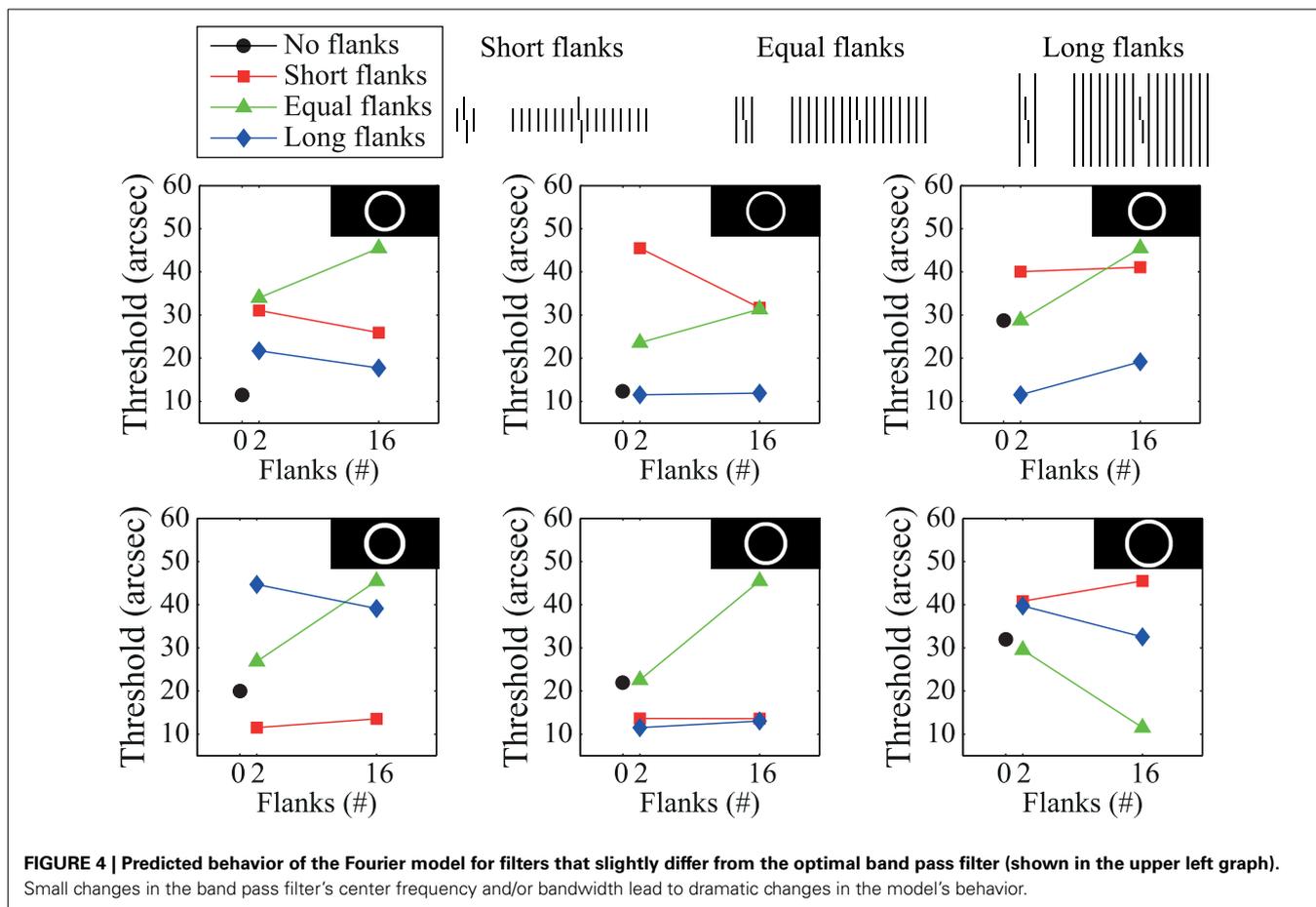
### 4. A FEEDBACK MODEL: WILSON-COWAN NEURAL NETWORK

We next considered a model that derives its key properties from the recurrent nature of information processing in a cooperative-competitive neural network. Variations of this kind of model have successfully accounted for visual masking data (Hermens et al., 2008) using stimuli very similar to those in **Figure 2**. The model first convolves the input image with an on-center, off-surround receptive field mimicking processing by the LGN. Next, the input activations are fed into both an excitatory and an inhibitory layer of neurons. Each layer convolves the input activations with a Gaussian blurring function and propagates activity



**FIGURE 3 | Simulation results using a Fourier model for the stimuli that produced the data presented in Figure 2.** Model results are plotted for representative low-pass filters (A,B), band-pass filters (C,D), and high-pass filters (E,F). Black and white insets show which frequencies were passed (white areas) and which frequencies were suppressed (black areas) in Fourier space (with lower frequencies in the center and higher frequencies near the

edges). The top row of subplots shows the best performance obtainable (using brute-force exhaustive search for the smallest sum of squared residuals against the human data) with each filter type. The bottom row shows results for filtering functions selected from different parts of the space - illustrating the variability of results obtainable with each filter type. The best overall performance we could obtain with this model is shown in (C).



over space with increasing time. The layers are reciprocally connected such that the excitatory units excite the inhibitory units and the inhibitory units inhibit the excitatory units. Details of the model, its filters, and its parameters can be found in Hermens et al. (2008) and Panis and Hermens (2014). Although the filters are local, the strength of activity at any given pixel location partly depends on the global pattern of activity across the network because of the feedback connections. When played out over time in a backward masking situation with stimuli similar to those in Figure 2, Hermens et al. (2008) showed that masking strength decreased as the number of flanking elements increased. More generally, the feedback in the network functions somewhat like a discontinuity detector by enhancing discontinuities and suppressing regularities. Panis and Hermens (2014) showed similar behavior for stimuli that produce crowding.

Since the model includes lateral feedback that promotes global processing, it satisfies the requirements identified above as “necessary” to explain crowding’s effects. Moreover, the model’s parameters were previously optimized for one stimulus, and then the model was validated by applying it to novel stimuli without further parameter optimization (Hermens et al., 2008). Thus, we would expect that any additional stimulus conditions that we apply this model to should require no further parameter optimization. We analyzed the model’s behavior in response to the stimuli used to generate the findings

in Figure 2 but found that the model performs poorly overall (Figure 5). In particular, the model produces virtually no difference between any of the conditions shown in Figure 5A. Figure 5B shows that the model also fails to reproduce the human data plotted in Figure 2B. Here, the model produces no substantial differences between the different flanker length conditions, it produces no crowding for the case where there are two flanking lines (thresholds are the same as in the unflanked case), and model thresholds always go up as an increasing function of the number of flankers (contrary to the human data).

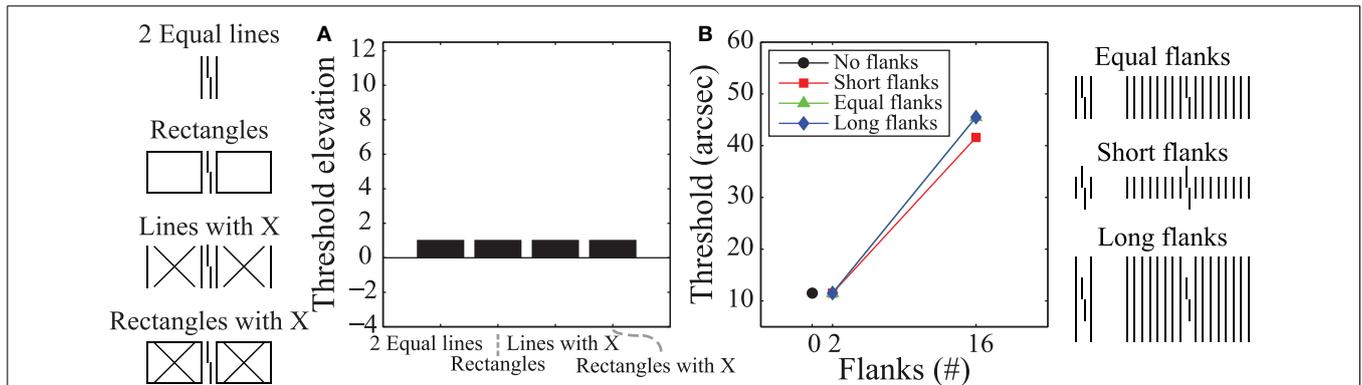
Even though the model has previously accounted for perceptual effects with similar kinds of stimuli and has strong feedback and global effects, the model simulations reported here do not account for the crowding effects in Figure 2. We cannot claim that the model architecture is fully rejected, as different filters and parameters may produce different model behaviors. Nevertheless, it is clear that global and feedback processing by themselves do not sufficiently constrain model properties relative to the observed crowding effects.

## 5. A FEEDBACK MODEL WITH PERCEPTUAL GROUPING: LAMINART NEURAL NETWORK

The previous simulations indicate that a model needs additional constraints beyond just feedback and global processing.

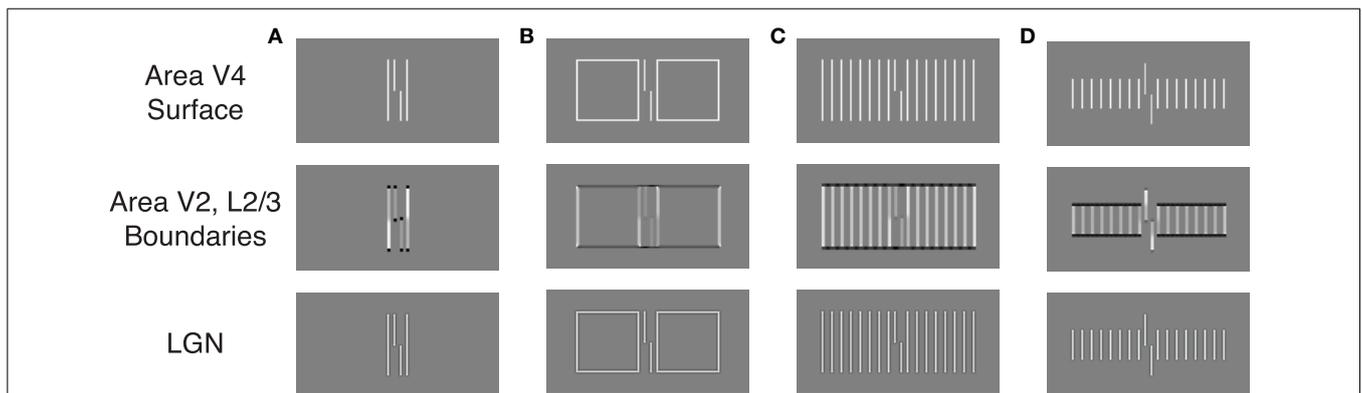
We next consider a model that has many additional constraints, the LAMINART model that has been proposed by Grossberg and colleagues (Raizada and Grossberg, 2001). The model is very complex and involves neural signals that interact across retinotopic coordinates, across laminar layers within a cortical area, and across cortical areas V1, V2, and V4. Various forms of the model account for neurophysiological and behavioral data related to depth perception (Grossberg, 1990; Grossberg and Howe, 2003), brightness perception (Grossberg and Todorovic, 1988), illusory contours (Grossberg and Mingolla, 1985), backward masking (Francis, 1997), and many other effects (Grossberg, 2003; Grossberg and Yazdanbakhsh, 2003). In particular, model simulations in Francis (2009) used stimuli very similar to those in Figure 2 to successfully account for a variety of backward masking effects. An integral part of the model explanations involved a form of perceptual grouping, which was indicated by the presence of illusory contours connecting elements within a group. Consistent with the ideas derived from Figure 2, these model grouping processes use feedback to generate global effects.

The model proposes separate processing streams for boundary and surface information. Grouping effects mostly occur in the boundary system through formation of illusory contours that connect nearly collinearly oriented edges, and Figures 6A,B show simulation results for two of the stimulus conditions in Figure 2A. When the flankers are two lines, the model generates boundary signals that represent each stimulus line (Area V2, Layer 2/3) and these boundary signals constrain brightness signals that pass from the LGN to Area V4. As a result, the Area V4 representation is essentially veridical relative to the original stimulus. As described in the Supplementary Material, the model signals are connected to human performance with a template matching process that tries to distinguish between verniers shifted to the left or right. Crowding effects occur because the vernier template (whose width is five times the spacing between stimulus elements) integrates information from both the flankers and the vernier target, thereby reducing the signal-to-noise ratio for vernier discrimination. In this way, the model matches the empirical finding that two flanking lines can produce crowding. Figure 7A shows



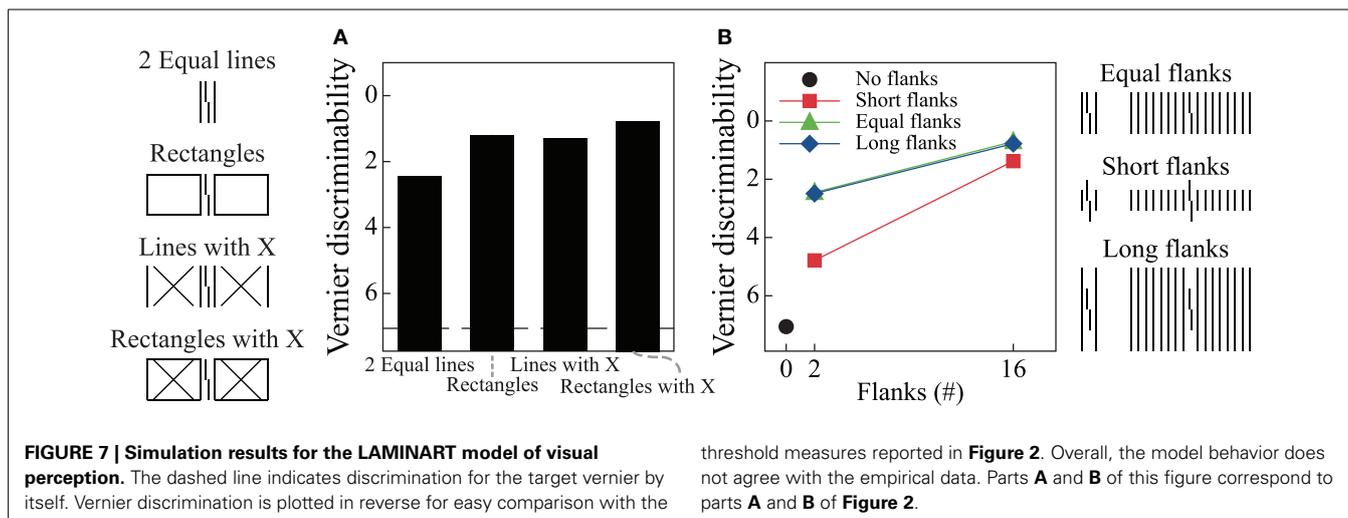
**FIGURE 5 | Results obtained using the model of Hermens et al. (2008) on the stimuli shown in Figure 2.** Compared with the human data plotted in Figure 2 this model does a poor job at capturing

human performance, despite using feedback signals that propagate information globally. Parts A and B of this figure correspond to parts A and B of Figure 2.



**FIGURE 6 | Simulation results for the LAMINART model of visual perception.** The bottom row indicates the activity of LGN cells and largely reflects the stimulus. The middle row schematizes the activity of orientation-sensitive neurons (dark gray to black indicates activity of a horizontally-tuned cell, light gray to white indicates

activity of a vertically tuned cell). The top row schematizes the activity of neurons that represent surfaces with perceived brightness. Judgments of target offset are based on the top row activities. (A–D) demonstrate the model’s behavior for several different types of flankers.



vernier discriminability (plotted in reverse for comparison with the threshold data) for this simulation, and it indicates that it is harder to identify a vernier with two flankers than to identify a vernier by itself (the dashed line).

**Figure 6B** shows the model's behavior when the flanking elements are rectangles. Although the local information is similar to that in the case of two flanking lines, the Area V2, Layer 2/3 cells respond quite differently by producing illusory contours that connect the two rectangles and the target vernier. Nevertheless, at the V4 filling-in stage, the perceptual representation is nearly veridical, and crowding occurs because the flanking elements again interfere with the vernier template matching calculations. Although the model has perceptual grouping, it incorrectly produces strong crowding where the empirical data indicate only weak crowding effects. These effects are indicated in **Figure 7A**, where the rectangle flankers condition indicates worse vernier discrimination than does the two equal-length flankers condition. Using flankers with an "X" produces the same pattern as for the conditions without an "X," and the rectangles provide the strongest masking. The data in **Figure 2A** shows the opposite pattern for the rectangles.

Similar properties exist for the stimuli producing the findings in **Figure 2B**. **Figures 6C,D** show the model's behavior in response to sixteen equal and short flankers. Consistent with the arguments about grouping described above, in the equal-length case the model generates illusory contours that connect the flankers with the target, thereby collectively grouping the flankers and target together. At the filling-in stage, all of the elements are represented and there is strong crowding. Also consistent with the above arguments, grouping is different for the short flankers (similar behavior would occur for the long flankers), such that the flanking elements are connected by illusory contours but the target remains separate. However, such grouping does not lead to a release from crowding in the model. At the Area V4 filling-in stage the flanking elements still interfere with the vernier discrimination process, even though the boundaries indicate that the flankers and target are part of different perceptual groups. **Figure 7B** shows that the LAMINART model does

not do a good job of matching the behavioral data in **Figure 2B**. This failure occurs even though the model includes feedback, has global effects, and contains grouping mechanisms that seem to operate much as recommended. Our claim is not that the model can be fully rejected by this failure, but we want to emphasize that a model with feedback, global processing, and mechanisms for perceptual grouping is not necessarily able to account for the observed human data.

## 6. WHAT CONSTRAINTS DOES A MODEL NEED?

The model simulations of crowding demonstrate that identification of global vs. local and feedback vs. feedforward processing does not necessarily promote the development of models that can account for human performance. We suspect the same kind of conclusion applies to models for many other visual phenomena. Although quantitative models of visual perception that account for visual processing often do include feedback and global processing (e.g., Bridgeman, 1971; Grossberg and Mingolla, 1985; Francis, 1997; Roelfsema, 2006; Craft et al., 2007; Kogo et al., 2010), this inclusion is often because such mechanisms provide specific computational properties that are needed to produce a functional visual system. The failure of the models discussed here relative to their success for other phenomena (e.g., backward masking) encourages a consideration of what kinds of constraints are useful for model development. It is unlikely that there is one single answer to this question, but we are willing to propose some ideas.

### 6.1. GLOBAL vs. LOCAL IS ABOUT INFORMATION REPRESENTATION

All models of visual processing involve encoding and representing information about the stimulus, and such a representation changes at various model stages so that some information is explicitly represented, other information is only implicitly represented, and some information is absent. A local model is one where the encoding of information about a certain position in visual space is modified only by information at nearby positions in space. In the case of crowding, the argument against local processing is that explicit or implicit information about the target

vernier appears to be affected by stimulus characteristics that are spatially far away in an unexpected way (e.g., two flanking squares produce less crowding than two flanking lines).

Even when the argument for non-local effects is convincing, it does not specify exactly how information about the target should be represented in a global-effects model. The crowding models described here include different types of information representation and different types of global effects. The Fourier model transforms spatial information into spectra and then applies a filtering step that loses some information about the target (as well as information about the flankers). The Wilson-Cowan model represents visual information in spatial (retinotopic) coordinates and introduces global effects via recurrent lateral inhibition. The LAMINART model also represents information in spatial coordinates, and it generates global effects via long-range illusory contours that connect spatially disparate boundaries, which can alter the boundary representations of the target. In practice, none of these global mechanisms produce crowding effects that emulate the behavioral data, at least in the instantiations considered here.

It seems to us that the global vs. local processing issue is something of a “red herring” that ignores deeper questions about the representation of visual information. A model must encode visual information in a way that allows for local or global processing, and identification of this encoding and its representation is the real model challenge. For example, in the LAMINART model, the information at the V4 surface stage provides a representation of information that (for the stimuli considered here) is essentially the same as the stimulus. Although there are groupings among boundaries, they do not modify the representation of visual information that is involved in the vernier offset judgment. What appears to be needed is for the boundary groupings to segment the visual information so that the target is represented separately from the flankers. In this way, the target’s offset could be discriminated with less interference from background elements. Francis (2009) described how such segmentation can occur for some visual masking situations that encode information about the target at the V4 surface representations in a different depth plane than information about the flankers. Such segmentation promotes good discrimination of the vernier offset. Foley et al. (2012) demonstrated that attentional effects could also produce similar segmentations in crowding conditions.

## 6.2. FEEDFORWARD vs. FEEDBACK IS ABOUT MODEL FUNCTION

Many of the discussions about feedforward vs. feedback processing seem predicated on the notion that if information is available at a model stage, then it can be used for a relevant task. For example, if binocular disparity information is available at V1, then it can be used for making depth discriminations at this stage. However, this attitude does not consider the many ways that feedback processing can influence information processing. In general, feedback processing tends to produce one of five robust model functions.

1. Completion: Excitatory feedback can “fill-in” missing information and thereby make explicit information that is implicitly represented by other aspects of an input pattern. One example of such completion is the generation of illusory contours in the LAMINART model, where the model explicitly represents “missing” contours that are justified by the co-occurrence of appropriate contours that are physically present. Another example of such completion is in the convergence of a Hopfield (1982) network to states with active neurons that were not directly excited by the input but are justified by their association with other active neurons.
2. Competition: A combination of excitatory and inhibitory feedback can enhance differences in neural activity and, in extreme forms, generate winner-take-all behavior in a network of neurons (Grossberg, 1973). Such networks can suppress noisy or irrelevant information and enhance the representation of other information in the system. For example, Wilson et al. (1992) proposed a competitive neural network to explain vector summation in motion perception, where units tuned to a particular motion direction inhibit units tuned to the orthogonal motion direction.
3. Preservation: Excitatory feedback can allow signals to persist well beyond the physical offset of a stimulus (e.g., Ögmen, 1993; Francis et al., 1994). Inhibitory feedback can also play a role in preservation of information by suppressing incoming signals that might alter the current pattern (Francis, 1997, 2000). A combination of excitatory and inhibitory linear feedback can also preserve pattern representations even with large changes in overall intensity (Grossberg, 1973). Grossberg et al. (2011), for example, used preservation to extend flickering stimuli long enough through time to allow their temporally integrated signals to generate smooth motion percepts.
4. Uniformization: Some types of non-linear feedback can diminish differences in neural activity and lead to uniform activity (Grossberg, 1973). Such information-losing feedback is not commonly used in neural models.
5. Comparison: Appropriately structured excitatory and inhibitory feedback can generate a signal that indicates the degree to which two neural activity patterns differ. Such signals are helpful for larger architectures that need to detect changes or stabilize learning (e.g., Grossberg, 1980; Sutton and Barto, 1998; Di Lollo et al., 2000).

These different functions often require rather different feedback mechanisms that involve the distribution of excitatory and inhibitory relations, the relative strength of feedback and feedforward signals, and the form of signal transformation between neurons. Thus, model development requires a characterization of function in order to be able to properly implement feedback. Characterizing model function is, of course, very challenging and generally requires some kind of over-arching theoretical framework to guide the computational goals of the model. For example, a model of crowding that theorizes a role for perceptual grouping needs to indicate how elements in a scene are identified as being “grouped,” explain the mechanisms by which such distinctions are generated, and characterize how such representations influence target processing and decision making. A focus on such functional details may reveal that a certain form of feedback processing is critical for the model to reproduce the human behavior (Raizada and Grossberg, 2001), or it may reveal that the

feedforward vs. feedback distinction is not as relevant as it first appeared (e.g., Francis and Hermens, 2002; Poder, 2013).

## 7. CONCLUSIONS

If the starting point of theorizing is that visual processing involves local interactions in a feedforward system, then it makes sense that investigations should explore whether such systems are sufficient to account for a given phenomenon. However, the modeling efforts presented here suggest that clear evidence of a role for global and feedback processing does not sufficiently constrain a model. At best, such investigations are only the starting point for model development, and further considerations are required concerning the details of information representation and model function. It might be easier to initiate theorizing by assuming global and feedback processing and then look for other more informative constraints such as task optimality, or perceptual completion.

## AUTHOR CONTRIBUTIONS

Aaron M. Clarke coded the simulations for the Fourier and Wilson-Cowan models. Gregory Francis coded the simulations for the LAMINART model. All authors contributed to the text.

## ACKNOWLEDGMENTS

Aaron Clarke was funded by the Swiss National Science Foundation (SNF) project “Basics of visual processing: what crowds in crowding?” (Project number: 320030\_135741). For Greg Francis, the research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 604102 (HBP).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fpsyg.2014.01193/abstract>

## REFERENCES

- Altmann, C. E., Bulthoff, H. H., and Kourtzi, Z. (2003). Perceptual organisation of local elements into global shapes in the human visual cortex. *Curr. Biol.* 13, 342–349. doi: 10.1016/S0960-9822(03)00052-6
- Atkinson, J., Anker, S., Evans, C., Hall, R., and Pimm-Smith, E. (1988). Visual acuity testing of young children with the cambridge crowding cards at 3 and 6 m. *Acta Ophthalmol.* 66, 505–508. doi: 10.1111/j.1755-3768.1988.tb04371.x
- Balas, B., Nakano, L., and Rosenholtz, R. (2009). A summary-statistic representation in peripheral vision explains visual crowding. *J. Vis.* 9, 13. doi: 10.1167/9.12.13
- Blakeslee, B., and McCourt, M. E. (1999). A multiscale spatial filtering account of the white effect, simultaneous brightness contrast and grating induction. *Vision Res.* 39, 4361–4377. doi: 10.1016/S0042-6989(99)00119-4
- Blakeslee, B., and McCourt, M. E. (2001). A multiscale spatial filtering account of the wertheimer-benary effect and the corrugated mondrian. *Vision Res.* 41, 2487–2502. doi: 10.1016/S0042-6989(01)00138-9
- Blakeslee, B., and McCourt, M. E. (2004). A unified theory of brightness contrast and assimilation incorporating oriented multiscale spatial filtering and contrast normalization. *Vision Res.* 44, 2483–2503. doi: 10.1016/j.visres.2004.05.015
- Blakeslee, B., Pasioka, W., and McCourt, M. E. (2005). Oriented multiscale spatial filtering and contrast normalization: a parsimonious model of brightness induction in a continuum of stimuli including White, Howe and simultaneous brightness contrast. *Vision Res.* 45, 607–615. doi: 10.1016/j.visres.2004.09.027
- Bridgeman, B. (1971). Metacontrast and lateral inhibition. *Psychol. Rev.* 78, 528–539. doi: 10.1037/h0031782
- Campbell, F. W., and Robson, J. G. (1968). Application of fourier analysis to the visibility of gratings. *J. Physiol.* 197, 551–566.
- Cesarei, A. D., and Loftus, G. R. (2011). Global and local vision in natural scene identification. *Psychon. Bull. Rev.* 18, 840–847. doi: 10.3758/s13423-011-0133-6
- Cope, D., Blakeslee, B., and McCourt, M. E. (2008). Structural theorems for simple cell receptive fields. *J. Vis.* 8:802. doi: 10.1167/8.6.802
- Cope, D., Blakeslee, B., and McCourt, M. E. (2009). Simple cell response properties imply receptive field structure: balanced gabor and/or bandlimited field functions. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* 26, 2067–2092. doi: 10.1364/JOSAA.26.002067
- Craft, E., Schuetze, H., Niebur, E., and von der Heydt, R. (2007). A neural model of figure-ground organization. *J. Neurophysiol.* 97, 4310–4326. doi: 10.1152/jn.00203.2007
- DeAngelis, G. C., Ohzawa, I., and Freeman, R. D. (1994). Length and width tuning of neurons in the cat's primary visual cortex. *J. Neurophysiol.* 71, 347–374.
- De Valois, R. L., Albrecht, D. G., and Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Res.* 22, 545–559. doi: 10.1016/0042-6989(82)90113-4
- Di Lollo, V., Enns, J. T., and Rensink, R. A. (2000). Competition for consciousness among visual events: the psychophysics of reentrant visual processes. *J. Exp. Psychol. Gen.* 129, 481–507. doi: 10.1037/0096-3445.129.4.481
- Dresp, B., and Grossberg, S. (1997). Contour integration across polarities and spatial gaps: from local contrast filtering to global grouping. *Vision Res.* 37, 913–924. doi: 10.1016/S0042-6989(96)00227-1
- Enns, J. T., and Di Lollo, V. (2000). What's new in visual masking? *Trends Cogn. Sci.* 4, 345–352. doi: 10.1016/S1364-6613(00)01520-5
- Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Am. A* 4, 2379–2394. doi: 10.1364/JOSAA.4.002379
- Foley, N., Grossberg, S., and Mingolla, E. (2012). Neural dynamics of object-based multifocal visual spatial attention and priming: object cueing, useful-field-of-view, and crowding. *Cogn. Psychol.* 65, 77–117. doi: 10.1016/j.cogpsych.2012.02.001
- Fourier, J. B. J. (1822). *La Théorie Analytique De La Chaleur*. Paris: Firmin Didot.
- Francis, G. (1997). Cortical dynamics of lateral inhibition: metacontrast masking. *Psychol. Rev.* 104, 572. doi: 10.1037/0033-295X.104.3.572
- Francis, G. (2000). Quantitative theories of metacontrast masking. *Psychol. Rev.* 107, 768. doi: 10.1037/0033-295X.107.4.768
- Francis, G. (2009). Cortical dynamics of figure-ground segmentation: shine-through. *Vision Res.* 49, 140–163. doi: 10.1016/j.visres.2008.10.002
- Francis, G., and Grossberg, S. (1995). *Cortical Dynamics of Boundary Segmentation and Reset: Persistence, Afterimages, and Residual Traces*. Technical Report, Boston University Center for Adaptive Systems and Department of Cognitive and Neural Systems.
- Francis, G., Grossberg, S., and Mingolla, E. (1994). Cortical dynamics of feature binding and reset: control of visual persistence. *Vision Res.* 34, 1089–1104. doi: 10.1016/0042-6989(94)90012-4
- Francis, G., and Hermens, F. (2002). Comment on: competition for consciousness among visual events: The psychophysics of reentrant visual processes, by di lollo, enns and rensink (2000). *J. Exp. Psychol. Gen.* 131, 590–593. doi: 10.1037/0096-3445.131.4.590
- Freeman, J., and Simoncelli, E. P. (2011). Metamers of the ventral stream. *Nat. Neurosci.* 14, 1195–1201. doi: 10.1038/nn.2889
- Freeman, R. D., Ohzawa, I., and Walker, G. (2001). Beyond the classical receptive field in the visual cortex. *Prog. Brain Res.* 134, 157–170. doi: 10.1016/S0079-6123(01)34012-8
- Gabor, D. (1946). Theory of communication. *J. Inst. Electr. Eng.* 93, 429–457.
- Greenwood, J. A., Bex, P. J., and Dakin, S. C. (2009). Positional averaging explains crowding with letter-like stimuli. *Proc. Natl. Acad. Sci. U.S.A.* 106, 13130–13135. doi: 10.1073/pnas.0901352106
- Grossberg, S. (1973). Contour enhancement, short term memory, and constancies in reverberating neural networks. *Stud. Appl. Math.* 52, 213–257.
- Grossberg, S. (1980). How does a brain build a cognitive code? *Psychol. Rev.* 87, 1–51. doi: 10.1037/0033-295X.87.1.1

- Grossberg, S. (1990). Neural facades: visual representations of static and moving form-and-color-and-depth. *Mind Lang.* 5, 411–456. doi: 10.1111/j.1468-0017.1990.tb00171.x
- Grossberg, S. (2003). Lamina cortical dynamics of visual form perception. *Neural Netw.* 16, 925–931. doi: 10.1016/S0893-6080(03)00097-2
- Grossberg, S., and Howe, P. D. (2003). A lamina cortical model of stereopsis and three-dimensional surface perception. *Vision Res.* 43, 801–829. doi: 10.1016/S0042-6989(03)00011-7
- Grossberg, S., Léveillé, J., and Versace, M. (2011). How do object reference frames and motion vector decomposition emerge in lamina cortical circuits? *Atten. Percept. Psychophys.* 73, 1147–1170. doi: 10.3758/s13414-011-0095-9
- Grossberg, S., and Mingolla, E. (1985). Neural dynamics of form perception: boundary completion, illusory figures, and neon color spreading. *Psychol. Rev.* 92, 173–211. doi: 10.1037/0033-295X.92.2.173
- Grossberg, S., and Rudd, M. E. (1989). A neural architecture for visual motion perception: group and element apparent motion. *Neural Netw.* 2, 421–450. doi: 10.1016/0893-6080(89)90042-7
- Grossberg, S., and Todorovic, D. (1988). Neural dynamics of 1-d and 2-d brightness perception: a unified model of classical and recent phenomena. *Percept. Psychophys.* 43, 241–277. doi: 10.3758/BF03207869
- Grossberg, S., and Yazdanbakhsh, A. (2003). Lamina cortical dynamics of 3-d surface stratification, transparency, and neon spreading. *J. Vis.* 3, 247. doi: 10.1167/3.9.247
- Harrison, L. M., Stephan, K. E., Rees, G., and Friston, K. J. (2007). Extra-classical receptive field effects measured in striate cortex with fmri. *Neuroimage* 34, 1199–1208. doi: 10.1016/j.neuroimage.2006.10.017
- Hermens, F., Luksys, G., Gerstner, W., Herzog, M. H., and Ernst, U. (2008). Modeling spatial and temporal aspects of visual backward masking. *Psychol. Rev.* 115, 83–100. doi: 10.1037/0033-295X.115.1.83
- Herzog, M. H., Ernst, U. A., Etzold, A., and Eurich, C. W. (2003). Local interactions in neural networks explain global effects in gestalt processing and masking. *Neural Comput.* 15, 2091–2113. doi: 10.1162/089976603322297304
- Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. U.S.A.* 79, 2554–2558. doi: 10.1073/pnas.79.8.2554
- Hubel, D. H., and Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.* 160, 106–154.
- Hung, C. P., Kreiman, G., Poggio, T., and DiCarlo, J. J. (2005). Fast readout of object identity from macaque inferior temporal cortex. *Science* 310, 863–866. doi: 10.1126/science.1117593
- Jones, J., and Palmer, L. (1987). An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex. *J. Neurophysiol.* 58, 1233–1258.
- Juan, C.-H., and Walsh, V. (2003). Feedback to V1: a reverse hierarchy in vision. *Exp. Brain Res.* 150, 259–263. doi: 10.1007/s00221-003-1478-5
- Keil, A., Sabatinelli, D., Ding, M., Lang, P. J., Ihssen, N., and Heim, S. (2009). Re-entrant projections modulate visual cortex in affective perception: evidence from granger causality analysis. *Hum. Brain Mapp.* 30, 532–540. doi: 10.1002/hbm.20521
- Keyser, C., and Perrett, D. I. (2002). Visual masking and {RSVP} reveal neural competition. *Trends Cogn. Sci.* 6, 120–125. doi: 10.1016/S1364-6613(00)01852-0
- Kogo, N., Strelcha, C., Gool, L. V., and Wagemans, J. (2010). Surface construction by a 2-d differentiation-integration process: a neurocomputational model for perceived border ownership, depth, and lightness in kanizsa figures. *Psychol. Rev.* 117, 406–439. doi: 10.1037/a0019076
- Lamme, V. A. F. (2006). Towards a true neural stance on consciousness. *Trends Cogn. Sci.* 10, 494–501. doi: 10.1016/j.tics.2006.09.001
- Lamme, V. A. F., and Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends Neurosci.* 23, 571–579. doi: 10.1016/S0166-2236(00)01657-X
- Levi, D. M., and Carney, T. (2009). Crowding in peripheral vision: why bigger is better. *Curr. Biol.* 19, 1988–1993. doi: 10.1016/j.cub.2009.09.056
- Levi, D. M., Klein, S. A., and Aitsebom, A. (1985). Vernier acuity, crowding and cortical magnification. *Vision Res.* 25, 963–977. doi: 10.1016/0042-6989(85)90207-X
- Livne, T., and Sagi, D. (2011). Multiple levels of orientation anisotropy in crowding with gabor flankers. *J. Vis.* 11, 1–10. doi: 10.1167/11.13.18
- Malania, M., Herzog, M. H., and Westheimer, G. (2007). Grouping of contextual elements that affect vernier thresholds. *J. Vis.* 7, 1–7. doi: 10.1167/7.2.1
- Manassi, M., Sayim, B., and Herzog, M. H. (2012). Grouping, pooling, and when bigger is better in visual crowding. *J. Vis.* 12, 1–14. doi: 10.1167/12.10.13
- Manassi, M., Sayim, B., and Herzog, M. H. (2013). When crowding of crowding leads to uncrowding. *J. Vis.* 13, 1–10. doi: 10.1167/13.13.10
- Öğmen, H. (1993). A neural theory of retino-cortical dynamics. *Neural Netw.* 6, 245–273. doi: 10.1016/0893-6080(93)90020-W
- Panis, S., and Hermens, F. (2014). Time course of spatial contextual interference: event history analysis of simultaneous masking by nonoverlapping patterns. *Am. Psychol. Assoc.* 40, 129–144. doi: 10.1037/a0032949
- Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., and Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nat. Neurosci.* 4, 739–744. doi: 10.1038/89532
- Poder, E. (2013). Attentional gating models of object substitution masking. *J. Exp. Psychol. Gen.* 142, 1130–1141. doi: 10.1037/a0030575
- Poggio, T., Mutch, J., Anselmi, F., Tacchetti, A., Rosasco, L., and Leibo, J. Z. (2013). “Does invariant recognition predict tuning of neurons in sensory cortex?” in *MIT: Computer Science and Artificial Intelligence Laboratory Technical Report* (Cambridge, MA), 1–12.
- Raizada, R. D., and Grossberg, S. (2001). Context-sensitive binding by the lamina circuits of V1 and V2: A unified model of perceptual grouping, attention, and orientation contrast. *Vis. Cogn.* 8, 431–466. doi: 10.1080/13506280143000070
- Rasche, C., and Koch, C. (2002). Recognizing the gist of a visual scene: possible perceptual and neural mechanisms. *Neurocomputing* 44–46, 979–984. doi: 10.1016/S0925-2312(02)00500-3
- Riesenhuber, M., and Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nat. Neurosci.* 2, 1019–1025.
- Roelfsema, P. R. (2006). Cortical algorithms for perceptual grouping. *Annu. Rev. Neurosci.* 29, 203–227. doi: 10.1146/annurev.neuro.29.051605.112939
- Sayim, B., Westheimer, G., and Herzog, M. H. (2010). Gestalt factors modulate basic spatial vision. *Psychol. Sci.* 21, 641–644. doi: 10.1177/0956797610368811
- Sceniak, M. P., Ringach, D. L., Hawken, M. J., and Shapley, R. (1999). Contrast's effect on spatial summation by macaque v1 neurons. *Nat. Neurosci.* 2, 733–739. doi: 10.1038/11197
- Schwartz, E. L. (1980). Computational anatomy and functional architecture of striate cortex: a spatial mapping approach to perceptual coding. *Vision Res.* 20, 645–669. doi: 10.1016/0042-6989(80)90090-5
- Serre, T., Oliva, A., and Poggio, T. (2007a). A feedforward architecture accounts for rapid categorization. *Proc. Natl. Acad. Sci. U.S.A.* 104, 6424–6429. doi: 10.1073/pnas.0700622104
- Serre, T., Wolf, L., Bileschi, S., Riesenhuber, M., and Poggio, T. (2007b). Robust object recognition with cortex-like mechanisms. *IEEE Trans. Pattern Anal. Mach. Intell.* 29, 411–426. doi: 10.1109/TPAMI.2007.56
- Serre, T., Wolf, L., and Poggio, T. (2005). “Object recognition with features inspired by visual cortex,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005* Vol. 2, (San Diego, CA), 994–1000.
- Sutton, R. S., and Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. Cambridge, MA: The MIT Press.
- Van den Berg, R., Roerdink, J. B., and Cornelissen, F. W. (2010). A neurophysiologically plausible population code model for feature integration explains visual crowding. *PLoS Comput. Biol.* 6:e1000646. doi: 10.1371/journal.pcbi.1000646
- VanRullen, R., Delorme, A., and Thorpe, S. J. (2001). Feed-forward contour integration in primary visual cortex based on asynchronous spike propagation. *Neurocomputing* 38–40, 1003–1009. doi: 10.1016/S0925-2312(01)00445-3
- Von der Heydt, R., Peterhans, E., and Baumgartner, G. (1984). Illusory contours and cortical neuron responses. *Science* 224, 1260–1262. doi: 10.1126/science.6539501
- Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Sing, M., et al. (2012). A century of gestalt psychology in visual perception: I. perceptual grouping and figure-ground organization. *Psychol. Bull.* 138, 1172–1217. doi: 10.1037/a0029333
- Wallis, T. S., and Bex, P. J. (2011). Visual crowding is correlated with awareness. *Curr. Biol.* 21, 254–258. doi: 10.1016/j.cub.2011.01.011
- Weisstein, E. (1968). A Rashevsky-Landahl neural net: Simulation of metacontrast. *Psychol. Rev.* 75, 494–521. doi: 10.1037/h0026769

- Whitney, D., and Levi, D. M. (2011). Visual crowding: a fundamental limit on conscious perception and object recognition. *Trends Cogn. Sci.* 15, 160–168. doi: 10.1016/j.tics.2011.02.005
- Wilkinson, F., Wilson, H. R., and Ellemberg, D. (1997). Lateral interaction in peripherally viewed texture arrays. *J. Optic. Soc. Am. A* 14, 2057–2068. doi: 10.1364/JOSAA.14.002057
- Wilson, H. R., Ferrera, V. P., and Yo, C. (1992). A psychophysically motivated model for two-dimensional motion perception. *Vis. Neurosci.* 9, 79–97. doi: 10.1017/S0952523800006386

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 20 June 2014; accepted: 02 October 2014; published online: 21 October 2014.*  
*Citation: Clarke AM, Herzog MH and Francis G (2014) Visual crowding illustrates the inadequacy of local vs. global and feedforward vs. feedback distinctions in modeling visual perception. Front. Psychol. 5:1193. doi: 10.3389/fpsyg.2014.01193*  
*This article was submitted to Perception Science, a section of the journal Frontiers in Psychology.*

*Copyright © 2014 Clarke, Herzog and Francis. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*