

Published in final edited form as:

Vision Res. 2014 November ; 0: 57–67. doi:10.1016/j.visres.2014.07.002.

Possible functions of contextual modulations and receptive field nonlinearities: pop-out and texture segmentation

Anita M. Schmid¹ and Jonathan D. Victor¹

Anita M. Schmid: ams2031@med.cornell.edu; Jonathan D. Victor: jdvicto@med.cornell.edu

¹Brain and Mind Research Institute, Division of Systems Neurology and Neuroscience, Weill Cornell Medical College, 1300 York Avenue, New York, NY, 10065, USA

Abstract

When analyzing a visual image, the brain has to achieve several goals quickly. One crucial goal is to rapidly detect parts of the visual scene that might be behaviorally relevant, while another one is to segment the image into objects, to enable an internal representation of the world. Both of these processes can be driven by local variations in any of several image attributes such as luminance, color, and texture. Here, focusing on texture defined by local orientation, we propose that the two processes are mediated by separate mechanisms that function in parallel. More specifically, differences in orientation can cause an object to “pop out” and attract visual attention, if its orientation differs from that of the surrounding objects. Differences in orientation can also signal a boundary between objects and therefore provide useful information for image segmentation. We propose that contextual response modulations in primary visual cortex (V1) are responsible for orientation pop-out, while a different kind of receptive field nonlinearity in secondary visual cortex (V2) is responsible for orientation-based texture segmentation. We review a recent experiment that led us to put forward this hypothesis along with other research literature relevant to this notion.

Keywords

visual cortex; boundary detection; cue invariance; spatial nonlinearities; surround suppression

1. Introduction

Our brain’s visual system has to solve several problems concurrently. One problem is choosing where to look at next, to make the best use of our limited area of highest acuity and limited attentional resources. Another problem is to segment the visual scene into objects, so that we can build an internal representation of the world around us and interact with it. Here,

© 2014 Elsevier Ltd. All rights reserved.

Corresponding Author: Anita Schmid Address: Brain and Mind Research Institute, Division of Systems Neurology and Neuroscience, Weill Cornell Medical College, 1300 York Avenue, New York, NY, 10065, USA, Phone: +1 (212) 746-6520, ams2031@med.cornell.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

we first outline the current thinking about these two processes, and how they are thought to be accomplished in the visual cortex. We next describe aspects of the response properties of single neurons in primary (V1) and secondary visual cortex (V2) and summarize how contextual modulations in V1 and receptive field properties in V2 are thought to be linked to the two processes. Then, we will summarize our recent neurophysiological study concluding that the two processes arise independently of each other in V1 and V2 respectively. Last, we will review psychophysical studies that are relevant to this notion and outline a general view of how these two processes are accomplished in the visual cortex through different non-linear interactions.

1.1 Pop-out

Visual search tasks are a powerful way of studying how the visual system directs attention. In visual search tasks, a target can be detected faster and with less effort if it differs in an elementary way from surrounding distractors (Treisman & Gelade, 1980). This is accompanied by a subjective impression of “pop-out” in that the target object seems to grab the viewer’s attention. Basic features for which a difference between target and distractors causes pop-out include orientation, color, motion, size and stereoscopic depth (see e.g. (Wolfe, 1994) for a review).

The pop-out phenomenon is linked to what is also called the “saliency” of a stimulus. Salient stimuli attract visual attention, and this is thought to be a way for the brain to decide which part of the visual environment to concentrate on, by either controlling eye-movements (overt attention) or by directing our visual attention without any associated eye-movement (covert attention). There is some controversy over what kinds of features influence the saliency of a stimulus and the specifics of this influence. For example, it is not clear in how far simple luminance contrast correlates with overt attention in humans (Einhauser & Konig, 2003; Frey, Konig, & Einhauser, 2007; Reinagel & Zador, 1999). The influence of color contrast on eye-movements, interestingly, depends on the image type (Frey, Honey, & Konig, 2008). However, texture contrasts are more consistently associated with eye-movement control (Frey et al., 2007; Krieger, Rentschler, Hauske, Schill, & Zetzsche, 2000; Parkhurst & Niebur, 2004).

Local texture contrast that is linked to saliency and pop-out could be detected by neurons in V1 via “contextual modulations”, a term describing the fact that a neuron’s responses to a stimulus within its receptive field can be modulated by stimuli outside of the receptive field. In this framework, the receptive field is defined as the region in visual space that can drive the neuron’s response on its own. This is sometimes also called the classical receptive field or receptive field center (Fitzpatrick, 2000). The modulatory influence of stimuli outside the receptive field, also called the non-classical receptive field or the receptive field surround, is found in many areas along the visual pathway and might serve as a means to make comparisons between stimuli inside and outside of the receptive field (Allman, Miezin, & McGuinness, 1985).

In V1, where many neurons are orientation-tuned, an important and well known kind of contextual modulation depends on the orientation of the stimulus presented in the surround. When the orientation of the stimulus in the center and the surround is the preferred

orientation, these neurons have a lower firing rate than when the orientation presented in the surround is different (Blakemore & Tobin, 1972; Fries, Albus, & Creutzfeldt, 1977; Nelson & Frost, 1978). This phenomenon, called “iso-orientation surround suppression,” is illustrated in Figure 1. In panel 1A, a grating of the preferred orientation is presented in the receptive field center (red circle); this elicits a high firing rate from the neuron. In panel 1B, only the receptive field surround (blue circle) is being stimulated, and the neuron does not respond (as expected, since the surround region – by definition – is modulatory and does not produce a response by itself). However, when the grating is presented to both the center and the surround, as shown in panel 1C, the neuron fires less than when only the center is being stimulated (compare to panel 1A); this is called “surround suppression”. Crucially, when the orientation of the grating in the surround is switched to an orientation perpendicular to the preferred orientation (panel 1D), the surround suppression is released, and the neuron fires more than when the surround grating also has the preferred orientation (compare to panel 1C).

Based on experiments in alert monkeys, Knierim and van Essen suggested that this type of contextual modulation might be the neurophysiological basis for orientation pop-out (Knierim & van Essen, 1992). Several neurophysiological studies in anesthetized cats and monkeys support this idea (Kastner, Nothdurft, & Pigarev, 1997, 1999; Nothdurft, Gallant, & Van Essen, 1999), as does a recent fMRI experiment in humans (Zhang, Zhaoping, Zhou, & Fang, 2012). Panels E and F in Figure 1 spell this out, by showing how iso-orientation surround suppression leads to higher firing rates in the pop-out condition. Panel 1E corresponds to the condition without pop-out – the target and distractor bars both have the preferred orientation – and because of the iso-orientation suppression, the neuron fires less. Panel 1F corresponds to the pop-out condition – the distractors have an orientation perpendicular to the target – and the iso-orientation surround suppression is released, leading to a higher firing rate.

Understanding the functional effects of surround suppression is complicated by the fact that the suppressive region in V1 is not always present all around the receptive field center. For example, the suppressive region can be confined to the “end-zones” along the length of the receptive field (also called “end-stopping”, “end-inhibition” or “length suppression”) or the “side flanks” of the receptive field (also called “side-inhibition” or “width suppression”) (DeAngelis, Freeman, & Ohzawa, 1994), or even oblique regions (Walker, Ohzawa, & Freeman, 1999). Figure 2 shows the impact of this asymmetry, using end-zone suppression as an example. Panel 2A shows how stimulating just the receptive field center with a rectangular patch of a sinusoidal grating in the preferred orientation activates the neuron. Panel 2B shows how presenting the same grating but only in the end-zones of the receptive field will not lead to any response (as expected, as in Fig. 1B, since these regions are only modulatory). In panel 2C, a rectangular patch with a sinusoidal grating of the preferred orientation covers the classical receptive field and also the end-zones, leading to a suppression of the response compared to when the classical receptive field is covered (compare to panel 2A). As shown in panel 2D, the suppression is released when patches with sinusoidal gratings of the orthogonal orientation are presented in the end-zones. Similarly, a stimulus consisting of several oriented bars of the preferred orientation will suppress the neuron if any of the bars align with the end-zones (panel 2E), and this

suppression will be released if the bars in the surround are of orthogonal orientation like in a classical pop-out stimulus (panel 2F).

The higher firing rate produced by iso-orientation surround suppression can also occur in situations without perceptual pop-out, because this mechanism produces an increased firing rate merely because of an orientation discontinuity. Panels 2G and 2H illustrate this: even just one line in the surround with the same orientation as in the center can suppress the neurons firing rate, if it happens to align with the suppressive zone (2G); when the orientation of the line in the surround is changed to the perpendicular orientation, part of that suppression is released (2H). Importantly, a two-bar display is not associated with perceptual pop-out, since neither token is distinguished from the background.

Given that neurons with iso-orientation surround suppression respond with higher firing rates to orientation discontinuities even when the stimulus is not a full-fledged orientation pop-out stimulus, it is no surprise that individual neurons in V1 cannot distinguish between pop-out stimuli and other stimuli that contain orientation discontinuities but don't elicit perceptual pop-out (Hegde & Felleman, 2003). Based on this observation, some studies suggest that orientation pop-out arises in higher visual areas such as V4 (Bogler, Bode, & Haynes, 2013; Burrows & Moore, 2009). Nevertheless, contextual effects in V1 neurons in anesthetized animals clearly display response properties that would allow the brain to extract the information necessary to detect orientation discontinuities and could therefore form the basis for orientation pop-out, even if the percept is only fully formed in higher visual areas.

1.2. Texture segmentation

Another process that can be driven by differences in orientations is texture segmentation. Sometimes an object boundary cannot be detected based on luminance cues or even color cues, but only based on texture cues. This is illustrated in Figure 3A: the elephant and the tree have similar color and luminance, but the texture difference of its skin compared to the tree provides a sufficient cue for the object boundary. In addition, detecting texture boundaries and color boundaries along with luminance boundaries is helpful for distinguishing shadows from object boundaries (Derrington et al., 2002; Johnson & Baker, 2004; Johnson, Kingdom, & Baker, 2005; Kingdom, 2003; Kingdom, Beauce, & Hunter, 2004; Schofield, Rock, Sun, Jiang, & Georgeson, 2010). This is because texture borders are often aligned with luminance borders in natural scenes (Johnson & Baker, 2004; Johnson et al., 2005), but shadows lying across an otherwise uniform surface are associated with luminance changes but no other cues such as color or texture changes. This is illustrated in Figure 3B: because there are confounding shadows on the tree trunk the visual system has to ignore the luminance boundaries that do not align with texture or color boundaries in order to accurately determine the shape of the trunk.

Luminance boundaries can be detected by simple linear filters, for example Gabor filters, which are often used as a simplified model of V1 receptive fields. Texture boundaries on the other hand, can only be detected by combining several such filters. Most proposed models consist of two stages of linear filtering with a rectifying nonlinearity in between, commonly called the filter-rectify-filter (FRF) model (Chubb & Sperling, 1988; Malik & Perona, 1990; Wilson, Ferrera, & Yo, 1992; Zavitz & Baker, 2013). The FRF model is illustrated in Figure

4, using a stimulus with a texture boundary between two regions with sinusoidal gratings of orthogonal orientation. In panel 4A, the filters in the first processing stage (bottom part of illustration) are oriented vertically, therefore only filters in the upper half of the stimulus are responding. Some of the individual filters will respond positively and others negatively, depending on the alignment of the filter and the luminance in the grating. At the rectification stage, negative responses are converted into positive responses, so that the sum of responses is now positive in the upper half of the stimulus and still zero in the lower half. In the second processing stage (top part of panel 4A), there is a filter of larger scale than the filters in the first processing stage, and it collects their rectified responses. Thus, it responds positively when the top part of the stimulus contains a vertical grating and the bottom part contains a grating that does not activate the filters in the first processing stage because of its contrasting orientation. The relationship between the orientation of filters in the first and second stage can be perpendicular to each other, as in panel 4A, but they can also have the same orientation, as illustrated in panel 4B. Here, only the bottom half of the same stimulus activates the linear filters in the first processing stage (bottom part of panel 4B). Thus, to ensure that the net response of the cell is still positive (as in panel A); the polarity of the second stage filter has to be reversed. If the orientations of the sinusoidal gratings were different, for example oblique, the filters in the first processing stage would have to have one of the orientations. Therefore, for each combination of two gratings that can signal a texture boundary, a different combination of first and second stage filters is needed.

The notion that texture boundary processing requires two stages of processing leads to the idea that the first stage could correspond to V1 neurons and the second stage to V2 neurons (Baker & Mareschal, 2001). Consistent with this idea are the several studies that have shown that neurons in monkey V2 and cat area 18 are selective for the texture boundaries (Baker et al., 2013; Leventhal, Wang, Schmolesky, & Zhou, 1998; Marcar, Raiguel, Xiao, & Orban, 2000; Mareschal & Baker, 1998; Zhou & Baker, 1994, 1996).

However, neurons in V1 also respond to texture boundaries, or more specifically orientation-defined texture boundaries, in a way that is consistent with their responses to pop-out stimuli or orientation discontinuities and also consistent with the contextual modulations described in Figures 1 and 2 (Nothdurft, Gallant, & Van Essen, 2000; Schmid, 2008). Figure 5 illustrates this for three different layouts of suppressive surround regions. In Figures 5A-C, the receptive field surround is isotropic, i.e., extends all around the receptive field center. For a continuous grating of the preferred orientation, the response is suppressed (panel 5A). Part of the suppression is released for an orientation-defined texture boundary; this release occurs both for a boundary that is orthogonal to the receptive field (panel 5B) and for a boundary that is parallel to the receptive field (panel 5C). In Figures 5D-F, the suppressive surround is confined to the end-zones of the receptive field. In this case, the suppression is released only when the orientation-defined texture boundary is orthogonal to the receptive field (panel 5E), but not when it is parallel (panel 5F). In Figures 5G-I, the suppressive surround is confined to the side-zones of the receptive field, and correspondingly, the suppression is not released when the orientation-defined texture boundary is orthogonal to the receptive field (panel 5H), but only when it is parallel (panel 5I).

Thus, both iso-orientation surround suppression (Figure 5) and the FRF model (Figure 4) can lead to increased responses when an orientation-defined texture boundary is presented.

One aspect of responses to texture boundaries that is helpful in distinguishing their possible roles in segmentation is “orientation cue invariance” – the notion that a neuron’s orientation tuning for standard (luminance) contours is the same as its orientation tuning for texture boundaries. This is a necessary characteristic of a neuron’s responses to texture-defined boundaries if it is to directly assist in identifying true object boundaries, since true object boundaries have both texture and luminance boundaries at the same orientation. However, for orientation-defined texture boundary responses due to iso-orientation surround suppression, orientation cue invariance is typically not present.

This is shown in Figure 5. If the suppression is all around the receptive field center (Fig. 5ABC) the orientation of the texture boundary does not matter. If the suppression is in the end-zones, the texture boundary needs to be orthogonal to the orientation tuning of the cell for luminance contours (Fig. 5DEC); if it is in the side-zones the texture boundary needs to be parallel to the receptive field. Therefore, depending on the structure of the suppression, the orientation tuning for the luminance boundaries and the texture boundaries may, or may not, be similar. Since the structure of suppression in individual neurons in V1 is very diverse (DeAngelis et al., 1994; Walker et al., 1999), there will be no consistency between the orientation tuning for luminance and texture boundaries across the V1 population (Tanaka & Ohzawa, 2009).

In primate V2 and cat area 18 however, a consistent alignment of orientation tuning between luminance and texture boundaries has been found (Leventhal et al., 1998; Mareschal & Baker, 1998; von der Heydt & Peterhans, 1989; Zhan & Baker, 2006). While the classical FRF model as shown in Figure 4 cannot account for this cue-invariance, an extension has been suggested that can. The basic FRF model generates a response to texture boundaries only and will not produce a response to luminance boundaries, unless very specific conditions apply (the subregions of the first-stage filter are not balanced, and the filters in the second stage have the same orientation as those in the first stage, like in Figure 4B, but with an asymmetric filter in the second stage). However, the FRF framework of Fig. 4 can be augmented by a parallel linear channel, as was proposed to explain the finding of cue-invariant orientation tuning between luminance borders and texture boundaries in V2 (Peterhans & von der Heydt, 1989). If this parallel channel is of similar shape as the filter in the second stage of the pure FRF model (Zhan & Baker, 2006), it will generate responses to ordinary gratings that have the same orientation tuning as the cell’s orientation tuning for texture boundaries.

1.3. Possible relationships between mechanisms underlying orientation based pop-out and texture segmentation

Because orientation-defined texture boundaries modulate responses in both V1 and V2 neurons, it is natural to ask how these responses are related to each other, and to contextual modulations. This is the focus of the remainder of the review.

One possibility is that V1 neurons respond to local orientation discontinuities independent of the orientation of the texture boundary, and that V2 neurons combine the output of those V1 neurons in a way that orientation cue-invariance is achieved (Heitger, von der Heydt, Peterhans, Rosenthaler, & Kübler, 1998; Nothdurft, 1991; Peterhans & von der Heydt, 1989; Schmid, 2008). This model is different from the FRF model in that the first processing stage is responsible for detection of local orientation discontinuities as well as standard luminance boundaries. Also, filtering is not required at the second processing stage, but rather, its role is to sum the outputs of the filters in the first stage along the axis of the receptive field of neuron in the second processing stage so that orientation cue-invariance is achieved. This summation of receptive fields is reminiscent of the classical Hubel & Wiesel model of how simple receptive fields in V1 could result from summation of LGN receptive fields in the appropriate spatial alignment (Hubel & Wiesel, 1962).

A second possibility is that responses to orientation-defined boundaries in V2 are based on contextual modulations in V2, formed independently of contextual modulations in V1. In this framework, V2 neurons combine the responses of V1 neurons to luminance boundaries to produce responses to orientation-defined boundaries, but do not rely on contextual modulations in V1 to detect orientation discontinuities. In this scenario, the mechanism by which V2 neurons respond to orientation-defined boundaries is iso-orientation surround suppression, just like observed in V1, just on a larger scale and with a bias for the surround suppression to come from the side-zones of the receptive field, so that orientation cue-invariance is produced.

These two possibilities make distinctive predictions. In the first case – that the signal generated by iso-orientation surround suppression in V1 is the starting point for the processing of orientation-defined boundaries in V2 neurons – manipulations that eliminate the V1 suppression should eliminate the orientation-defined boundary responses in V2. We found a condition in which the surround suppression in V1 is eliminated, but, as we show in the next section, the V2 response to orientation-defined boundaries persists, therefore we can exclude this possibility.

The second possibility – that responses to orientation-defined boundaries in V2 are generated independently of contextual modulations in V1, but also by surround suppression – predicts that texture boundary responses and surround suppression in V2 should be correlated on a neuron-by-neuron basis. Moreover, if the responses to boundaries in V2 are orientation cue-invariant, meaning that the preferred orientation is the same for luminance and orientation-defined boundaries, then these responses should specifically correlate with suppression localized in the side-zones of the receptive field (see Figure 5). These predictions are also tested in the next section, and, as we show, they do not hold.

Therefore, we conclude that responses to orientation-defined boundaries in V2 are generated independently of contextual modulations in V1 and also they are not based on surround suppression in V2. Instead, they must be based on another kind of nonlinearity, such as the FRF model. Because the FRF model does not fully account for the observed response properties of V2 neurons – as we will detail below – we propose an extension of it, which adds a second stage of rectification.

The hypothesis that we put forward here is that while contextual modulations (iso-orientation surround suppression) in V1 signal local orientation discontinuities that may underlie perceptual orientation pop-out, orientation-defined boundary responses in V2 arise independently and could be the basis for texture segmentation.

2. Detailed spatiotemporal analysis of receptive field interactions in V1 and V2

An experiment that tests the hypothesis that contextual modulation signals arising in V1 are necessary for orientation-defined boundary signals in V2 must meet two criteria. First, the same type of stimulus must be used for assaying neurons in both regions. Second, there needs to be a way to manipulate the stimulus so that the contextual signal in V1 is lost. Further, an experiment that tests the hypothesis that neurons in V2 have orientation cue-invariant responses to luminance-defined and orientation-defined boundaries needs to probe at least two orientations for both types of boundaries.

Figure 6 shows the paradigm we used that meets these criteria (Schmid, Purpura, & Victor, 2014). The stimulus (Figure 6A) consisted of a grid of adjacent rectangular regions covering the receptive field center and surround. Each rectangular region was filled with a sinusoidal grating of either the preferred orientation or the orthogonal non-preferred orientation, and for each stimulus frame the orientation within each region was assigned pseudo-randomly using a binary sequence. The stimulus frame duration was either 20 ms or 40 ms.

Because the orientation of each region was assigned in a pseudorandom fashion on each frame, we could use a reverse correlation technique to separate several different contributions to the response (Figure 6B). By correlating the response of the neuron to the orientation in each of the regions, we obtained a measurement of how much the orientation of the luminance boundary within each region influences the neurons firing rate. In parallel, by correlating the response of the neuron to whether the orientation in a pair of neighboring regions was either the same or different, we obtained a response to the local orientation-defined texture boundary. Importantly, there are two different possible orientations of texture boundaries, namely parallel to receptive field (blue lines in Figure 6A) as well as orthogonal to the receptive field (green lines in Figure 6A). By comparing responses to these two kinds of texture boundaries, we were also able to see if neurons preferred one orientation of texture boundary over the other.

The population summary of the responses of neurons to luminance and texture boundaries at the two frame durations is summarized in Figure 7. We found that in V1 responses to the luminance boundaries are present both at 20 ms and 40 ms frame durations (black traces in panels A and D), but the responses to orientation-defined boundaries arose only at 40 ms frame duration (green and blue traces in panels E and F). There were also negative responses to luminance boundaries (red traces in panel D), consistent with what one would expect given iso-orientation surround suppression, as described in the Introduction and illustrated in Figure 5. Specifically, across the population, these responses bore no fixed relationship to the response to the luminance boundary: some neurons responded to such boundaries if they were orthogonal to the receptive fields, while others responded to them if they were parallel

to the receptive field (compare panels E and F). Thus, at the population level, V1 neurons did not exhibit orientation cue-invariance.

V2 neurons also had responses to the orientation-defined texture boundaries in this paradigm, but these differed in two key ways from the V1 responses. First, they were present for the 20 ms frame duration (blue traces in panel I), a condition in which V1 texture boundary responses were absent (compare with panel C). This shows that the first hypothesis mentioned in the Introduction cannot be true; the response to orientation-defined boundaries in V2 cannot rely on the V1 orientation-defined boundary responses. The second difference between the orientation-defined texture boundary responses in V2 and V1 was that in V2, strong responses to orientation-defined boundaries only occurred if they were parallel to the receptive field (blue traces in panels I and L, but only one green trace in panels H and K). Hence, V2 neurons that respond to texture boundaries are orientation cue-invariant.

To understand the implications of different dynamics in V1 and V2, it is important to distinguish between temporal resolution and latency. The fact that texture boundary responses in V2 were present at shorter frame durations than in V1 does not mean that these signals are faster in latency (i.e., appear sooner) in V2 than in V1 – just that visual signals are processed with finer temporal resolution in V2 than in V1. In fact, our data is consistent with the expected flow of information from V1 to V2: the response latency for luminance boundaries is not shorter in V2 than in V1 (compare starting time of black traces in all panels of Figure 7). However the responses to texture boundaries in V2 have a higher temporal resolution than those in V1, i.e. they appear with stimulus frame durations of only 20 ms. This means that the texture boundary response in V2 arise independently of texture boundary responses in V1 but still may rely on the local luminance boundary information that flows from V1 to V2.

Since V2 texture boundary responses arise independently of those in V1, we next considered the mechanisms by which they might be generated: specifically, whether the V2 responses arose from iso-orientation surround suppression, vs. an alternative.

To test the hypothesis that the response to orientation-defined texture boundaries in V2 is based on iso-orientation surround suppression, we measured surround suppression in the same neurons studied with the paradigm of Figure 6, using standard grating stimuli. V1 responses served as a benchmark. Consistent with the notion that V1 responses to orientation-defined boundaries reflected iso-orientation surround suppression, we found that these two aspects of responses were correlated. V1 neurons that responded to orientation-defined boundaries orthogonal to the receptive field tended to have suppression in the end-zones (length suppression), while those that responded to orientation-defined boundaries parallel to the receptive field tended to have suppression in the side-zones (width suppression). This is what one expects, if the responses to orientation-defined boundaries in V1 are based on iso-orientation surround suppression.

Neurons in V2, on the other hand, showed no correlation between responses to orientation-defined boundaries and surround suppression. Even though surround suppression was

present in V2 neurons and comparable in extent to surround suppression in V1, the V2 neurons that responded to orientation-defined boundaries were no more likely to have surround suppression than the V2 neurons that did not respond to such boundaries. In addition, neurons that responded to orientation-defined boundaries parallel to the receptive field and happened to have surround suppression did not necessarily have the surround suppression in the side-zones. And among the neurons with suppression in the end-zones, none responded to orientation-defined boundaries orthogonal to the receptive field.

In sum, the responses to orientation-defined boundaries in V1 are consistent with iso-orientation surround suppression, but not so in V2. Hence, only the third hypothesis remains viable: V2 texture boundary responses arise independently from those in V1, and are generated by a mechanism other than surround suppression.

The observed responses of V2 neurons to texture boundaries is largely consistent with the filter-rectify-filter model, along with the addition of a second luminance channel as discussed in Section 1.3. The second parallel channel ensures that the neuron also responds to luminance boundaries and not only to texture boundaries, as we found. However, we found another property of the texture boundary responses in V2 that require an additional extension of the FRF model. For example the FRF models shown in Figure 4 will only respond if the vertical grating is in the top half of the stimulus and the horizontal grating is in the bottom half; it will not respond at all if the orientations of the two gratings are reversed. What we find in V2 neurons of anesthetized macaques is different: the neurons will respond similarly to both texture border polarities. This response property can be incorporated into the FRF model by adding a second rectification stage after the second filtering stage, resulting in a filter-rectify-filter-rectify (FRFR) cascade. In this framework, the first filtering and rectification produces “complex cells” that are non-selective for the spatial phase of the luminance boundaries and the second filtering and rectification stage produces cells that are non-selective for the spatial phase of texture boundaries. Interestingly, the need to augment the FRF model by a second stage of rectification to account for processing of texture-defined contours was also noted in a psychophysical study of perception of high-order spatial correlations (Victor & Conte, 1991).

3. Two perceptual processes

In our recent neurophysiological experiment discussed in Section 2 we showed that frame durations that are too short to generate contextual modulations in V1 nevertheless produce nonlinear interactions in V2. We also proposed that the contextual modulations in V1 are the basis of pop-out, while the receptive field nonlinearities in V2 are the basis of orientation-defined texture segmentation. Together, these notions predict that the dynamics of pop-out are more sluggish than that of texture segregation. While this has not directly been tested in a psychophysical experiment, there are two studies that bear on this question, which we now review.

3.1. Orientation Pop-Out

We hypothesize that orientation pop-out is based on contextual modulations in V1 in macaques as well as humans. Given that -- in anesthetized macaque -- luminance boundary

responses in V1 are present at 20 ms as well as 40 ms frame duration, but contextual modulations in V1 are only present at 40 ms frame duration, we predict that human perception of a oriented bar has a higher temporal resolution than the perception of an orientation pop-out stimulus.

This was directly tested by the study of Nothdurft (2000), which concentrated on the dynamics of pop-out. Nothdurft's Experiment 2 compared detection of an isolated line with detection of the same line in an orientation pop-out paradigm. The key variable was stimulus presentation time (10 – 150 ms) and the stimulus was immediately followed by a mask consisting of crosses. The presentation time needed for the criterion performance was always longer for the orientation pop-out than for the single line (30 ms vs 20 ms for 75% performance, averaged across subjects; see Figure 7a in Nothdurft, 2000). A potential confound is that the detection of the single line does not necessarily have to depend on the activation of orientation tuned neurons in V1, as activation of neurons based on luminance differences would suffice. But this concern was eliminated by a follow-up experiment that showed that the identification of the orientation of a single line (and not just detecting its presence) also had a higher temporal resolution than the detection of orientation pop-out, in the sense that the same performance was achieved with shorter stimulus presentation times (Nothdurft, 2002).

Since latency and temporal resolution are independent measures of response dynamics, analysis of latency might provide another way to link neural responses with behavior. Several studies have shown that the latency of contextual effects in V1 is longer than the latency of the response to the single elements of the stimulus (Bair, Cavanaugh, & Movshon, 2003; Knierim & van Essen, 1992; Nothdurft et al., 1999). We confirmed this finding in our recent experiment as well (Schmid et al., 2014). It has been suggested that longer latency carries the expectation that longer presentation times are needed for perception (Knierim & van Essen, 1992; Nothdurft, 2000), but the direct correlate of neural response latency is reaction time. Even so, a comparison between neural response latency and reaction time is hindered by many factors, such as stimulus parameters, experiment conditions, decision process, and motor times, which are non-trivial to model (Miller & Ulrich, 2003). Although several researchers have used reaction time measurements to test the hypothesis that V1 is involved in computing the saliency of pop-out stimuli (Koene & Zhaoping, 2007; Zhaoping & Zhe, 2012), there does not appear to be a study that compared the reaction time for identification of a single oriented line to the reaction time for this identification in the context of orientation pop-out stimulus.

3.2. Orientation-defined Texture Segmentation

Another set of studies (Motoyoshi & Nishida, 2001) provided measurements of the dynamics of orientation-defined texture boundaries, and, below we describe how the findings relate to our proposal. To describe the logic of their experiment, it is helpful to view the detection of texture boundaries as a two-stage process. In particular, for orientation-defined texture segmentation, the first stage consists of orientation coding, and the second stage extracts orientation contrast. This view parallels the filter-rectify-filter model described in the Introduction, where the first stage consist of filters detecting luminance

boundaries at a given orientation, and the second stage consists of filters detecting differences in texture. We assume that the first stage corresponds to the responses of V1 neurons to luminance-defined boundaries, which are present at a stimulus frame duration of 20 ms in anesthetized macaques. Our hypothesis is then that the second stage corresponds to the responses to orientation-defined texture boundaries in V2, which are also present at a stimulus frame duration of 20 ms. Therefore, we predict that the temporal resolution cut-off for the orientation coding stage and the orientation-contrast coding stage are similar.

To assess the dynamics of these stages psychophysically, Motoyoshi and Nishida (Motoyoshi & Nishida, 2001) used a special kind of oriented patch – the so-called D2 pattern – as the texture element. D2 patterns have the useful property that when a D2 pattern at one orientation is superimposed on a D2 pattern in the orthogonal orientation, the result is a circular Laplacian (target-shaped) pattern, which has no orientation cues at all. Motoyoshi & Nishida exploited this property by constructing their stimuli from alternating pairs of texture images, in which the orientation of each local D2 pattern is orthogonal on the two frames. If the alternation frequency is higher than the resolution of the orientation coding stage, only circular patches are perceived.

To measure the temporal frequency resolution of both processing stages (the orientation coding stage and the orientation contrast coding stage), they compared perception of two kinds of alternating-D2 texture images. In the first, the orientations of the patches in the target region differed by 90 degrees from those in the background. In this condition, the perception of the target region is limited by the orientation-contrast coding stage: the stimulus frame duration must be long enough to perceive the local orientation and also long enough to compute the orientation difference. In the second, the orientations in the target region differed by 45 degrees from those in the background. In this condition, the perception of the target region is limited only by the orientation coding stage: if the stimulus frame duration is long enough to perceive the local orientation information, the system can integrate over several frames to detect the target region, because the target and the background activate different sets of orientation detectors.

Their logic is the following: if the temporal resolution of the orientation coding stage is higher than that of the subsequent stage of orientation contrast coding, the upper temporal frequency limit for the 45 degree texture should be higher than for the 90 degree texture. On the other hand, if the orientation coding is the temporal bottleneck, the upper temporal frequency limit of both textures should be similar. The latter is the alternative that our hypothesis predicts: since the orientation difference signal in V2 neurons is already present at 20 ms, we hypothesize that the limiting factor is the temporal resolution of the orientation coding in V1 (also present at 20 ms) and no extra integration time is needed to compute the orientation difference. This is what they find: the temporal frequency limit for the 45 degree texture is not higher than for the 90 degree texture

We note that while the findings of Motoyoshi and Nishida (2001) are in line with our prediction that segmentation based on orientation contrast is as rapid as orientation coding itself, there is a discrepancy between the temporal resolutions that they measured, and our physiologic measurements. The upper temporal limit measured psychophysically ranged

from 11 – 16 Hz in two subjects (71% performance in a four alternative forced choice; see Figure 6 in Motoyoshi and Nishida, 2001); the corresponding time in each stimulus phase (approximately 30 to 45 ms) is longer than what we found in anesthetized macaques (20 ms). There are many obvious differences between the experiments that could account for this discrepancy, including species, stimulus characteristics, and the choice of threshold criterion.

3.3. Transition between pop-out and texture segmentation

The psychophysical studies discussed above, along with our recent neurophysiological study, suggest that the temporal resolution for orientation pop-out and orientation-defined texture segmentation is different. Some stimuli seem to activate only the more sluggish orientation pop-out mechanisms, other stimuli recruit the faster orientation-defined texture segmentation. The study of Motoyoshi and Nishida was focused on texture segmentation, but their systematic variation of spatial parameters (experiment 3) included conditions that resembled classic orientation pop-out stimuli. We hypothesize that orientation pop-out relies on V1 contextual modulations, which are slower than simple orientation coding in V1. Therefore, for stimulus conditions which drive only the pop-out process and not the texture segmentation process, the upper temporal frequency limit for the 45 degree texture (limited by orientation coding) should be higher than for the 90 degree texture (limited by orientation-contrast coding). This is in fact what they found for some special cases, in which the target region consisted only of one oriented patch (see their figures 13 and 14). While the authors hypothesized that this reversal of the upper temporal frequency limit for the 45 and 90 degrees texture was due to the fact that in those conditions the texture segmentation was very difficult, we propose a more specific reason; in those conditions – resembling orientation pop-out rather than texture segmentation – the faster orientation-contrast coding stage in V2 is not activated and therefore the more sluggish signal in V1 has to be used.

4. Conclusions

Pop-out and texture segmentation are distinct processes. One directs attention and the other is important for parsing an image into objects. Based on the studies presented here, we hypothesize that contextual modulations in V1 drive orientation pop-out, and nonlinear receptive field interactions in V2 drive texture segmentation based on orientation differences. The contextual modulations in V1 are consistent with iso-orientation surround suppression, but the receptive field non-linearities in V2 can only be described by a more complex model, for example a filter-rectify-filter-rectify model with a second parallel luminance channel. Since the connections in this FRFR model need to be very specific, it remains to be seen how such a complex response behavior might be achieved in biologically plausible neural network model. We also conclude that those two processes function in parallel, and the properties of the stimulus determine whether perception is dominated by one process or the other.

Acknowledgments

This work was funded by NIH grant EY9314 to J.D. Victor.

References

- Allman J, Miezin F, McGuinness E. Stimulus specific responses from beyond the classical receptive field: neurophysiological mechanisms for local-global comparisons in visual neurons. *Annu Rev Neurosci.* 1985; 8:407–430. [PubMed: 3885829]
- Bair W, Cavanaugh JR, Movshon JA. Time course and time-distance relationships for surround suppression in macaque V1 neurons. *J Neurosci.* 2003; 23(20):7690–7701. [PubMed: 12930809]
- Baker C, Li G, Wang Z, Yao Z, Yuan N, Talebi V, Zhou Y. Second-Order Neuronal Responses to Contrast Modulation Stimuli in Primate Visual Cortex. *J Vis.* 2013
- Baker CL Jr, Mareschal I. Processing of second-order stimuli in the visual cortex. *Prog Brain Res.* 2001; 134:171–191. [PubMed: 11702543]
- Blakemore C, Tobin EA. Lateral inhibition between orientation detectors in the cat's visual cortex. *Exp Brain Res.* 1972; 15(4):439–440. [PubMed: 5079475]
- Bogler C, Bode S, Haynes JD. Orientation pop-out processing in human visual cortex. *Neuroimage.* 2013; 81:73–80. [PubMed: 23689014]
- Burrows BE, Moore T. Influence and limitations of popout in the selection of salient visual stimuli by area V4 neurons. *J Neurosci.* 2009; 29(48):15169–15177. [PubMed: 19955369]
- Chubb C, Sperling G. Drift-balanced random stimuli: a general basis for studying non-Fourier motion perception. *J Opt Soc Am A.* 1988; 5(11):1986–2007. [PubMed: 3210090]
- DeAngelis GC, Freeman RD, Ohzawa I. Length and width tuning of neurons in the cat's primary visual cortex. *J Neurophysiol.* 1994; 71(1):347–374. [PubMed: 8158236]
- Derrington AM, Parker A, Barraclough NE, Easton A, Goodson GR, Parker KS, Webb BS. The uses of colour vision: behavioural and physiological distinctiveness of colour stimuli. *Philos Trans R Soc Lond B Biol Sci.* 2002; 357(1424):975–985. [PubMed: 12217169]
- Einhauser W, König P. Does luminance-contrast contribute to a saliency map for overt visual attention? *Eur J Neurosci.* 2003; 17(5):1089–1097. [PubMed: 12653985]
- Fitzpatrick D. Seeing beyond the receptive field in primary visual cortex. *Curr Opin Neurobiol.* 2000; 10(4):438–443. [PubMed: 10981611]
- Frey HP, Honey C, König P. What's color got to do with it? The influence of color on visual attention in different categories. *J Vis.* 2008; 8(14):6 1–17. [PubMed: 19146307]
- Frey HP, König P, Einhauser W. The role of first- and second-order stimulus features for human overt attention. *Percept Psychophys.* 2007; 69(2):153–161. [PubMed: 17557586]
- Fries W, Albus K, Creutzfeldt OD. Effects of interacting visual patterns on single cell responses in cats striate cortex. *Vision Res.* 1977; 17(9):1001–1008. [PubMed: 595408]
- Hegde J, Felleman DJ. How selective are V1 cells for pop-out stimuli? *J Neurosci.* 2003; 23(31):9968–9980. [PubMed: 14602810]
- Heitger F, von der Heydt R, Peterhans E, Rosenthaler L, Kübler O. Simulation of neural contour mechanisms: representing anomalous contours. *Image and Vision Computing Computational and psychophysical studies of early vision.* 1998; 16(6–7):407–421.
- Hubel DH, Wiesel TN. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol.* 1962; 160:106–154. [PubMed: 14449617]
- Johnson AP, Baker CL Jr. First- and second-order information in natural images: a filter-based approach to image statistics. *J Opt Soc Am A Opt Image Sci Vis.* 2004; 21(6):913–925. [PubMed: 15191171]
- Johnson AP, Kingdom FA, Baker CL Jr. Spatiochromatic statistics of natural scenes: first- and second-order information and their correlational structure. *J Opt Soc Am A Opt Image Sci Vis.* 2005; 22(10):2050–2059. [PubMed: 16277276]
- Kastner S, Nothdurft HC, Pigarev IN. Neuronal correlates of pop-out in cat striate cortex. *Vision Res.* 1997; 37(4):371–376. [PubMed: 9156167]
- Kastner S, Nothdurft HC, Pigarev IN. Neuronal responses to orientation and motion contrast in cat striate cortex. *Vis Neurosci.* 1999; 16(3):587–600. [PubMed: 10349977]
- Kingdom FA. Color brings relief to human vision. *Nat Neurosci.* 2003; 6(6):641–644. [PubMed: 12740582]

- Kingdom FA, Beauce C, Hunter L. Colour vision brings clarity to shadows. *Perception*. 2004; 33(8): 907–914. [PubMed: 15521690]
- Knierim JJ, van Essen DC. Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *J Neurophysiol*. 1992; 67(4):961–980. [PubMed: 1588394]
- Koene AR, Zhaoping L. Feature-specific interactions in saliency from combined feature contrasts: evidence for a bottom-up saliency map in V1. *J Vis*. 2007; 7(7):6 1–14. [PubMed: 17685802]
- Krieger G, Rentschler I, Hauske G, Schill K, Zetzsche C. Object and scene analysis by saccadic eye-movements: an investigation with higher-order statistics. *Spat Vis*. 2000; 13(2–3):201–214. [PubMed: 11198232]
- Leventhal AG, Wang Y, Schmolesky MT, Zhou Y. Neural correlates of boundary perception. *Vis Neurosci*. 1998; 15(6):1107–1118. [PubMed: 9839975]
- Malik J, Perona P. Preattentive texture discrimination with early vision mechanisms. *J Opt Soc Am A*. 1990; 7(5):923–932. [PubMed: 2338600]
- Marcas VL, Raiguel SE, Xiao D, Orban GA. Processing of kinetically defined boundaries in areas V1 and V2 of the macaque monkey. *J Neurophysiol*. 2000; 84(6):2786–2798. [PubMed: 11110809]
- Mareschal I, Baker CL Jr. A cortical locus for the processing of contrast-defined contours. *Nat Neurosci*. 1998; 1(2):150–154. [PubMed: 10195131]
- Miller J, Ulrich R. Simple reaction time and statistical facilitation: a parallel grains model. *Cogn Psychol*. 2003; 46(2):101–151. [PubMed: 12643892]
- Motoyoshi I, Nishida S. Temporal resolution of orientation-based texture segregation. *Vision Res*. 2001; 41(16):2089–2105. [PubMed: 11403793]
- Nelson JJ, Frost BJ. Orientation-selective inhibition from beyond the classic visual receptive field. *Brain Res*. 1978; 139(2):359–365. [PubMed: 624064]
- Nothdurft H. Saliency from feature contrast: temporal properties of saliency mechanisms. *Vision Res*. 2000; 40(18):2421–2435. [PubMed: 10915883]
- Nothdurft HC. Texture segmentation and pop-out from orientation contrast. *Vision Res*. 1991; 31(6): 1073–1078. [PubMed: 1858322]
- Nothdurft HC. Latency effects in orientation popout. *Vision Res*. 2002; 42(19):2259–2277. [PubMed: 12220582]
- Nothdurft HC, Gallant JL, Van Essen DC. Response modulation by texture surround in primate area V1: correlates of "popout" under anesthesia. *Vis Neurosci*. 1999; 16(1):15–34. [PubMed: 10022475]
- Nothdurft HC, Gallant JL, Van Essen DC. Response profiles to texture border patterns in area V1. *Vis Neurosci*. 2000; 17(3):421–436. [PubMed: 10910109]
- Parkhurst DJ, Niebur E. Texture contrast attracts overt visual attention in natural scenes. *Eur J Neurosci*. 2004; 19(3):783–789. [PubMed: 14984430]
- Peterhans E, von der Heydt R. Mechanisms of contour perception in monkey visual cortex. II. Contours bridging gaps. *J Neurosci*. 1989; 9(5):1749–1763. [PubMed: 2723748]
- Reinagel P, Zador AM. Natural scene statistics at the centre of gaze. *Network-Computation in Neural Systems*. 1999; 10(4):341–350.
- Schmid AM. The processing of feature discontinuities for different cue types in primary visual cortex. *Brain Res*. 2008; 1238:59–74. [PubMed: 18771659]
- Schmid AM, Purpura KP, Victor JD. Responses to orientation discontinuities in v1 and v2: physiological dissociations and functional implications. *J Neurosci*. 2014; 34(10):3559–3578. [PubMed: 24599456]
- Schofield AJ, Rock PB, Sun P, Jiang X, Georgeson MA. What is second-order vision for? Discriminating illumination versus material changes. *J Vis*. 2010; 10(9):2. [PubMed: 20884600]
- Tanaka H, Ohzawa I. Surround suppression of V1 neurons mediates orientation-based representation of high-order visual features. *J Neurophysiol*. 2009; 101(3):1444–1462. [PubMed: 19109456]
- Treisman AM, Gelade G. A feature-integration theory of attention. *Cogn Psychol*. 1980; 12(1):97–136. [PubMed: 7351125]
- Victor JD, Conte MM. Spatial organization of nonlinear interactions in form perception. *Vision Res*. 1991; 31(9):1457–1488. [PubMed: 1949617]

- von der Heydt R, Peterhans E. Mechanisms of contour perception in monkey visual cortex. I. Lines of pattern discontinuity. *J Neurosci*. 1989; 9(5):1731–1748. [PubMed: 2723747]
- Walker GA, Ohzawa I, Freeman RD. Asymmetric suppression outside the classical receptive field of the visual cortex. *J Neurosci*. 1999; 19(23):10536–10553. [PubMed: 10575050]
- Wilson HR, Ferrera VP, Yo C. A psychophysically motivated model for two-dimensional motion perception. *Vis Neurosci*. 1992; 9(1):79–97. [PubMed: 1633129]
- Wolfe JM. Guided Search 2.0 A revised model of visual search. *Psychon Bull Rev*. 1994; 1(2):202–238. [PubMed: 24203471]
- Zavitz E, Baker CL. Texture sparseness, but not local phase structure, impairs second-order segmentation. *Vision Research*. 2013; 91(0):45–55. doi: <http://dx.doi.org/10.1016/j.visres.2013.07.018>. [PubMed: 23942289]
- Zhan CA, Baker CL Jr. Boundary cue invariance in cortical orientation maps. *Cereb Cortex*. 2006; 16(6):896–906. [PubMed: 16151176]
- Zhang X, Zhaoping L, Zhou T, Fang F. Neural activities in v1 create a bottom-up saliency map. *Neuron*. 2012; 73(1):183–192. [PubMed: 22243756]
- Zhaoping L, Zhe L. Properties of V1 neurons tuned to conjunctions of visual features: application of the V1 saliency hypothesis to visual search behavior. *PLoS One*. 2012; 7(6):e36223. [PubMed: 22719829]
- Zhou YX, Baker CL Jr. Envelope-responsive neurons in areas 17 and 18 of cat. *J Neurophysiol*. 1994; 72(5):2134–2150. [PubMed: 7884449]
- Zhou YX, Baker CL Jr. Spatial properties of envelope-responsive cells in area 17 and 18 neurons of the cat. *J Neurophysiol*. 1996; 75(3):1038–1050. [PubMed: 8867116]

- Orientation pop-out and texture segmentation are separate and function in parallel.
- Orientation pop-out is linked to contextual modulations in V1.
- These contextual effects are consistent with iso-orientation surround suppression.
- Texture segmentation is linked to receptive field interactions in V2.
- These response properties are consistent with filter-rectify-filter-rectify models.

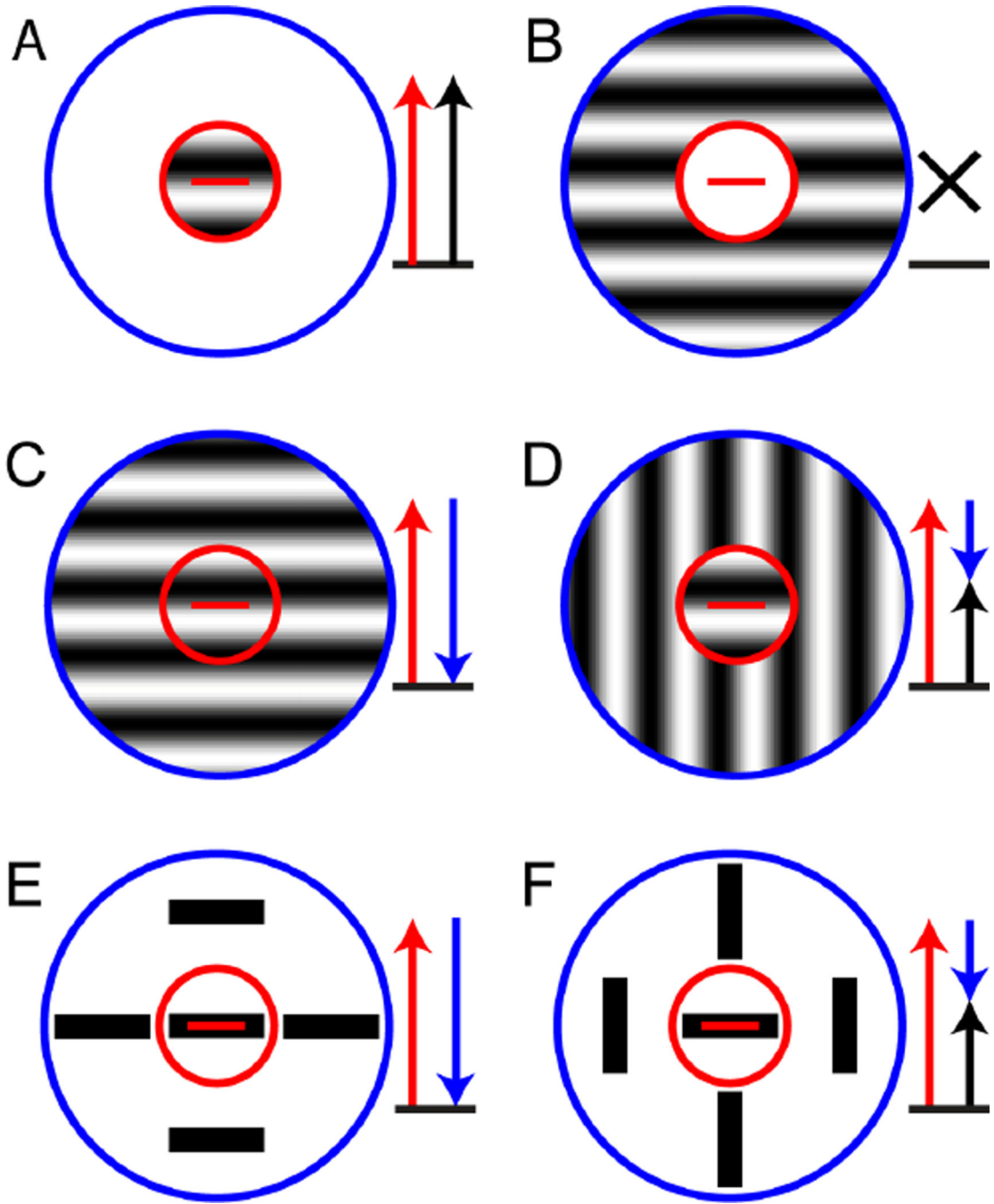


Figure 1. Iso-orientation surround suppression and pop-out

These panels diagram the responses of an idealized cortical neuron with a receptive field center (red circle) that is tuned to the horizontal orientation (red bar) and an iso-oriented suppressive surround (blue circle). Red arrows represent the excitation due to stimulation of the receptive field center, blue arrows represent suppression due to stimulation of the surround and black arrows represent the net response. (A) A grating of the preferred orientation confined to the receptive field center elicits a high firing rate (red arrow). (B) A grating of the preferred orientation confined to the suppressive surround elicits no response

(black cross). (C) A grating of the preferred orientation covering both the center and the surround elicits a small response, because the center contribution (red arrow) is reduced by the suppressive surround (blue arrow). (D) When the grating in the surround is switched to a non-preferred orientation, part of the suppression is released (short blue arrow) and the overall firing rate is high. (E and F) As in C and D, but with the gratings are replaced by bars, typically used in pop-out experiments.

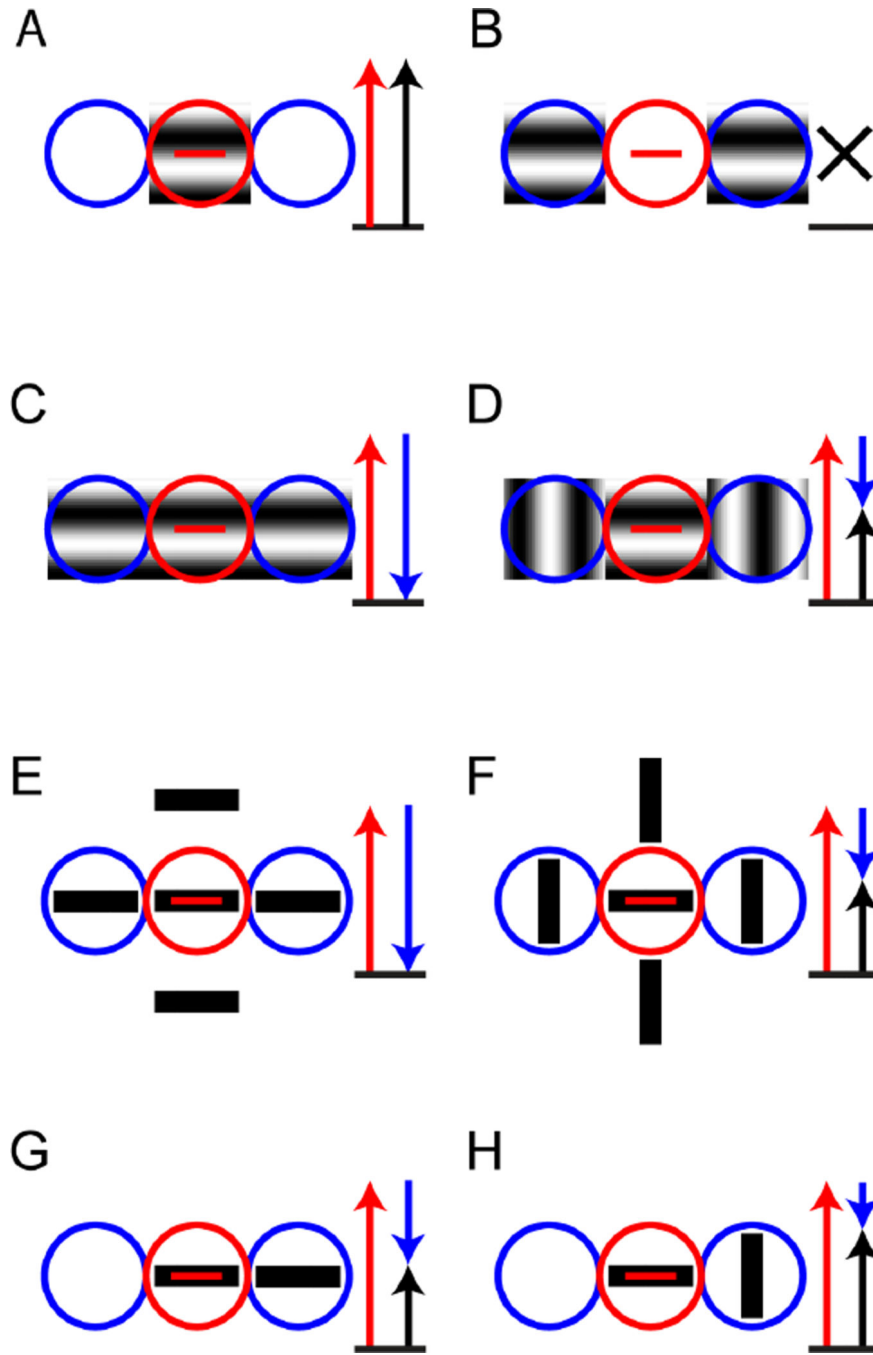


Figure 2. Iso-orientation surround suppression in the end-zones

These panels diagram the responses of an idealized cortical neuron with a receptive field center (red circle) that is tuned to the horizontal orientation (red bar) and iso-oriented suppression confined to the end-zones (blue circles). Red arrows represent the excitation due to stimulation of the receptive field center, blue arrows represent suppression due to stimulation of the surround and black arrows represent the net response. (A) A grating of the preferred orientation confined to the receptive field center elicits a high firing rate (red arrow). (B) A grating of the preferred orientation confined to the end-zones elicits no

response (black cross). (C) A grating of the preferred orientation covering both the center and the end-zones elicits a small response, because the center contribution (red arrow) is reduced by the suppressive end-zones (blue arrow). (D) When the grating in the end-zones is switched to a non-preferred orientation, part of the suppression is released (short blue arrow) and the overall firing rate is high. (E) A stimulus consisting of bars of preferred orientation in center and surround, typical of pop-out experiments, elicits a small response due to end-zone suppression. (F) When the surrounding bars have an orientation orthogonal to the preferred orientation, suppression is released. (G) A stimulus consisting of two bars of preferred orientation, one in center and one in one of the end-zones, produces a modest response because of partial end-zone suppression. (H) When the bar in the end-zone has the orthogonal orientation, the partial suppression is released and the response is larger.

A



B



Figure 3. Texture boundaries in natural images

(A) The boundary between the elephant and the tree is visible primarily because of a difference in texture. (B) In this image, object boundaries are defined by differences in luminance, color, and texture, while shadow boundaries are defined only by differences in luminance.

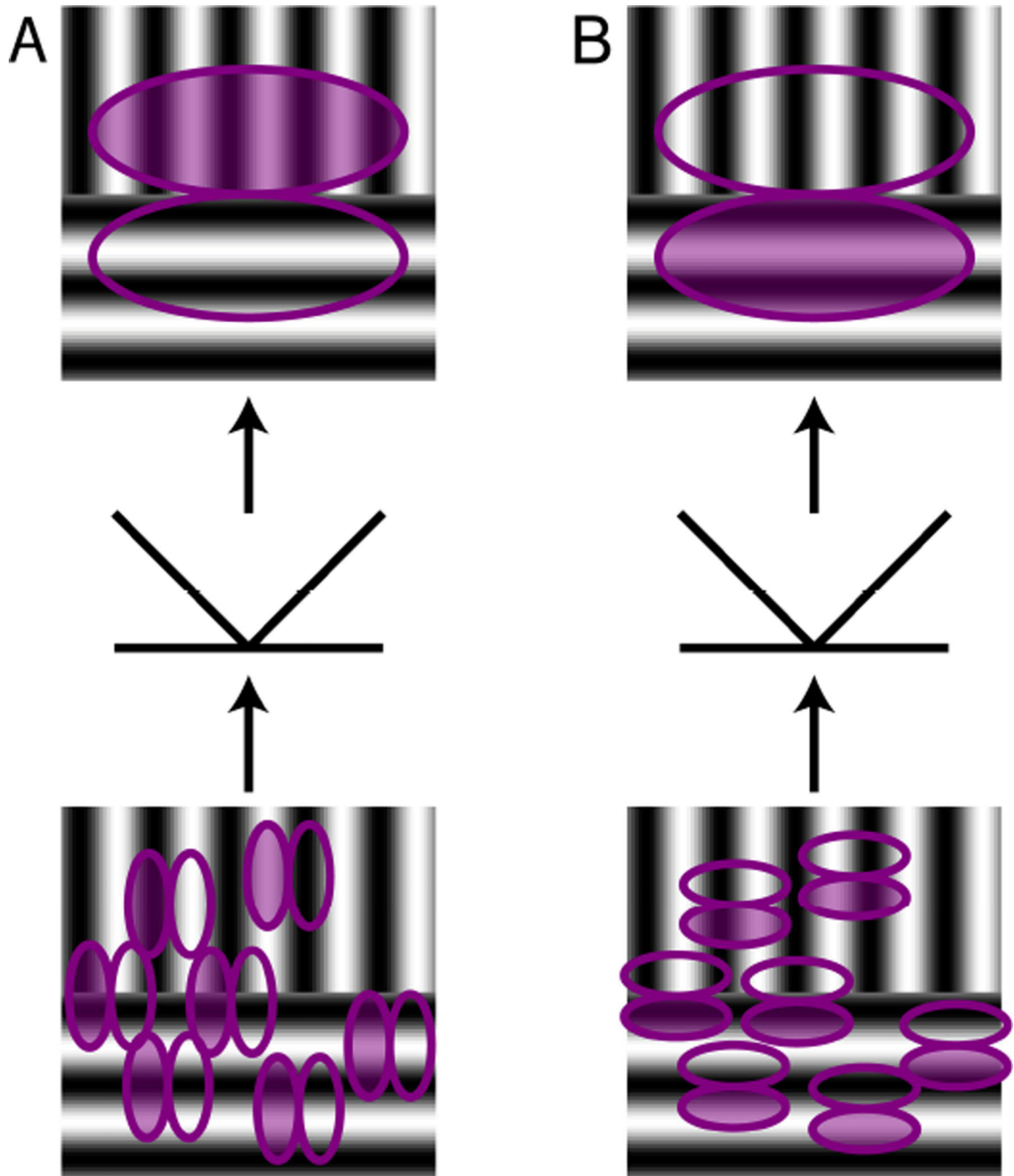


Figure 4. Filter-Rectify-Filter model

Panels A and B show two variants of a model consisting of two filtering stages with an intervening rectification stage. The first stage, shown in the bottom part of each panel, consists of a population of oriented linear filters (filled ellipse for positive component, unfilled ellipse for negative component) that act on stimulus luminance. The outputs of these individual filters are then rectified. The second stage, shown in the top part of each panel, consists of a filter acting on the rectified outputs of the first stage, resulting in an orientation difference filter (filled ellipse for positive component, unfilled for negative component). (A)

If the first-stage filters have an orientation preference orthogonal to second-stage filter, a positive response is generated if the positive component of a horizontal second-stage filter overlays the vertical grating while the negative component overlays the horizontal grating. (B) If the first-stage filters have the same orientation preference as the second-stage filter, a positive response is generated if the positive component of a horizontal second-stage filter overlays the horizontal grating while the negative component overlays the vertical grating. Note that in both A and B, the second stage generates a negative response if the two gratings that define the contour are interchanged.

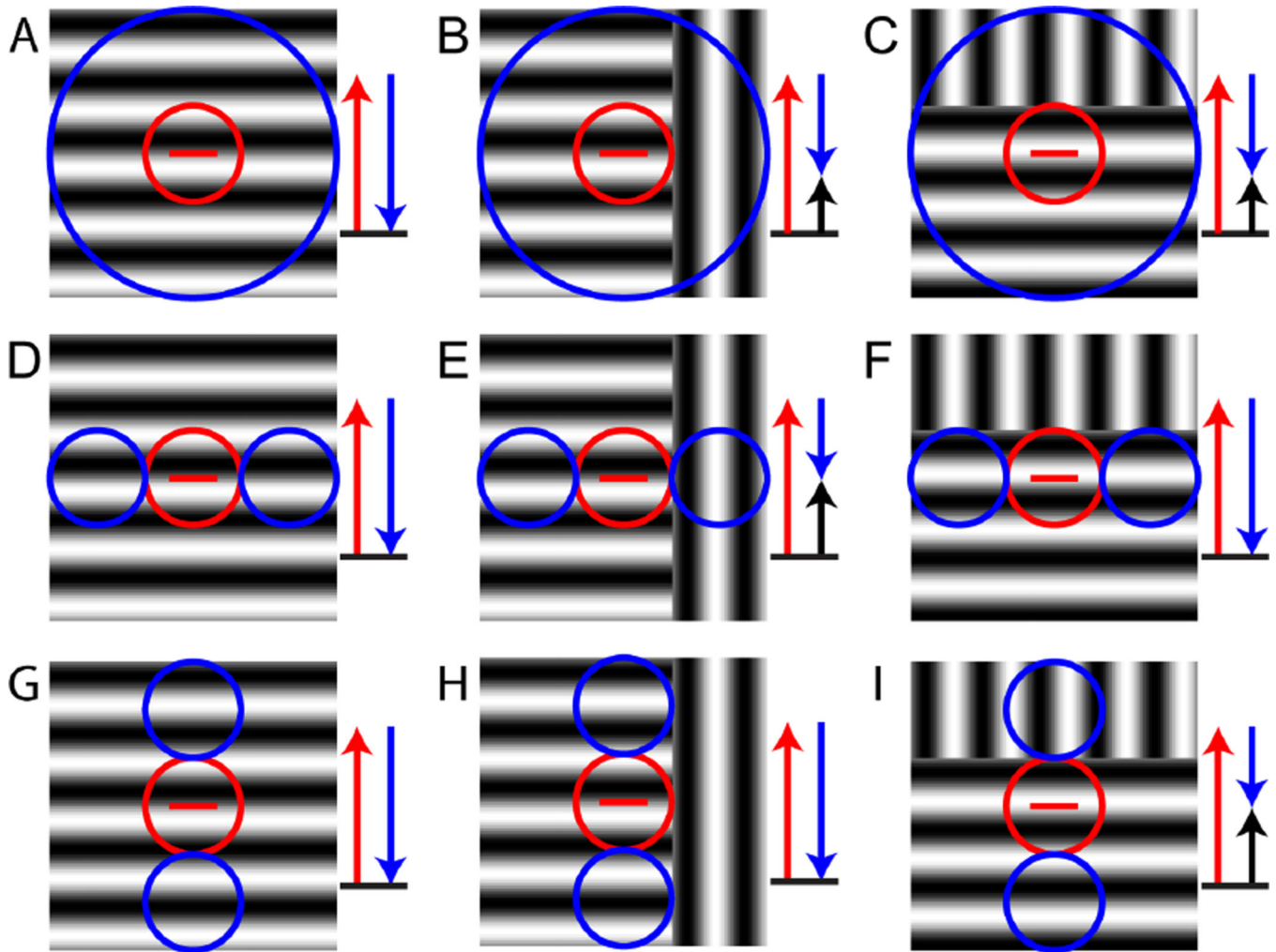


Figure 5. Iso-orientation surround suppression and texture segmentation

These panels diagram the responses of three idealized cortical neurons (each row) to three kinds of grating stimuli (each column). The three neurons have a receptive field center (red circle) that is tuned to the horizontal orientation (red bar) and are distinguished by the layouts of their iso-oriented suppressive surrounds (blue circles). The stimuli consist of a full-field grating (left column) and orientation-defined texture boundaries orthogonal to the preferred orientation (middle column) or parallel to it (right column). Red arrows represent the excitation due to stimulation of the receptive field center, blue arrows represent suppression due to stimulation of the surround and black arrows represent the net response. For all stimuli, the neuron responds because the receptive field center is covered by a grating of the preferred orientation (red arrow). Top row: circumferential suppressive surround. The full-field grating in the preferred orientation (A) activates surround suppression (blue arrow) and reduces the response. (B and C) Suppression is partially released when an orientation-defined texture boundary is present, because a portion of the surround is covered by the orientation orthogonal to the preferred orientation. This occurs both for a texture boundary orthogonal to the preferred orientation (B) and parallel to it (C). Second row: suppressive surround confined to the end-zones. Suppression by a full-field grating (D) is released when

the orientation-defined boundary is orthogonal to the preferred orientation of the center (E) but not when it is orthogonal to it (F). Third row: suppressive surround confined to the side-zones. Suppression by a full-field grating (G) is not released when the orientation-defined boundary is orthogonal to the preferred orientation of the center (H), but is released when the boundary is parallel to it (I).

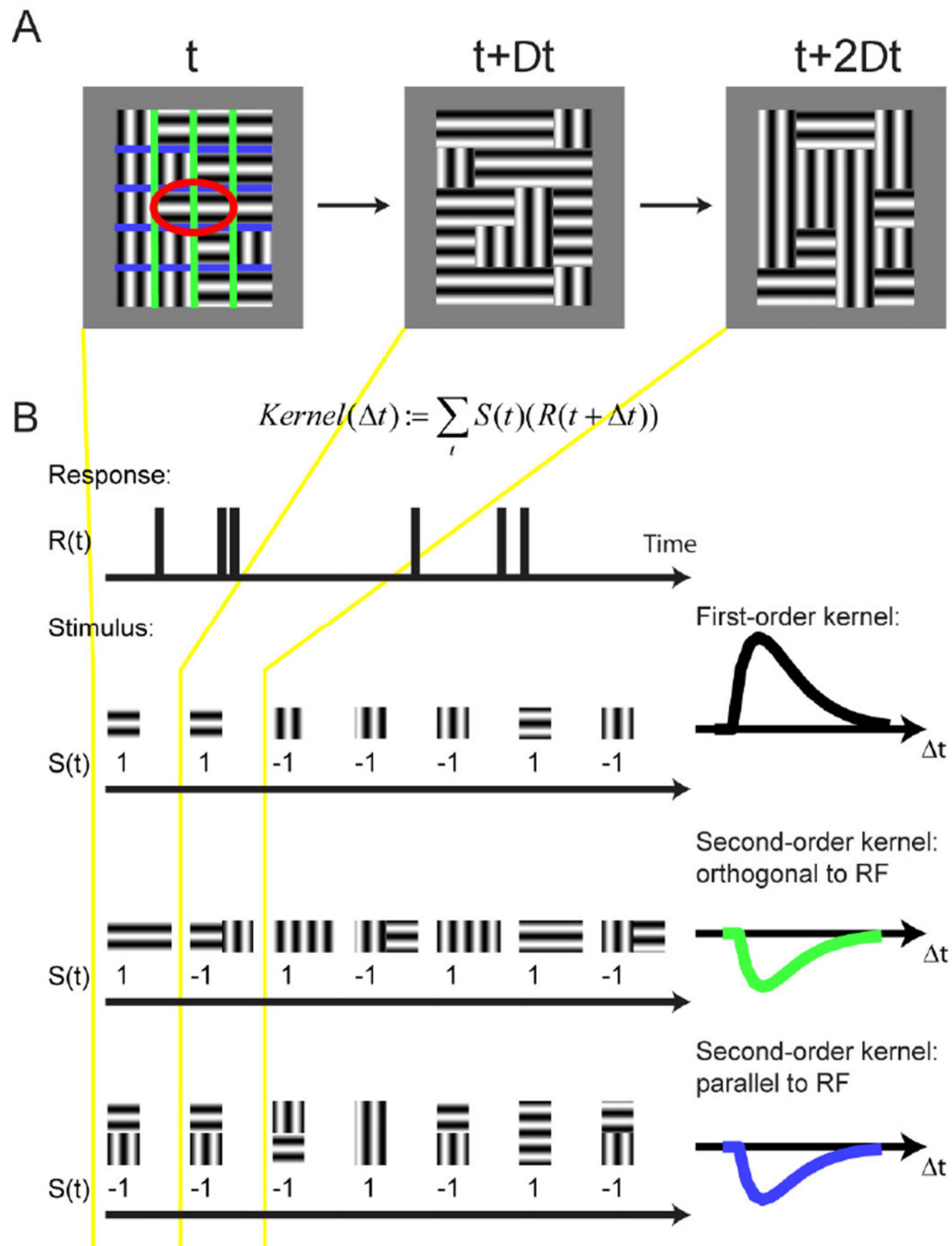


Figure 6. Orientation-discontinuity stimulus

(A) A grid of rectangular regions covered the classical receptive field (red ellipse) and surrounding space. The stimulus was aligned with the preferred orientation of the receptive field. Each region contained a static sinusoidal grating, either in the preferred or the orthogonal, non-preferred orientation. The orientation in each region was randomly reassigned at every time step, which was either 20 or 40 ms. Colored lines show the region boundaries orthogonal (green) and parallel (blue) to the receptive field; these lines were not part of the stimulus. (B) Computation of first- and second-order kernels. For each region in

the stimulus, the neuron's spike response, in 10 ms bins, was cross-correlated with the stimulus sequence, coded as +1 for the preferred orientation and -1 for the orthogonal orientation. Cross-correlation of this sequence against the neural response yielded the first-order kernel in that region. For the computation of each spatial second-order kernel, the response was correlated with the product of the values of the stimulus in the two neighboring regions: 1 if the grating orientation in the two regions was equal and -1 if they were different. Therefore, positive kernel values indicate an enhanced response to texture continuity, and negative values indicate an enhanced response to texture discontinuity. Note that we distinguished two kinds of spatial second-order kernels: one for regions that shared an edge orthogonal to the receptive field's long axis, and one for regions that shared an edge that was parallel to it. Adapted from (Schmid et al., 2014).

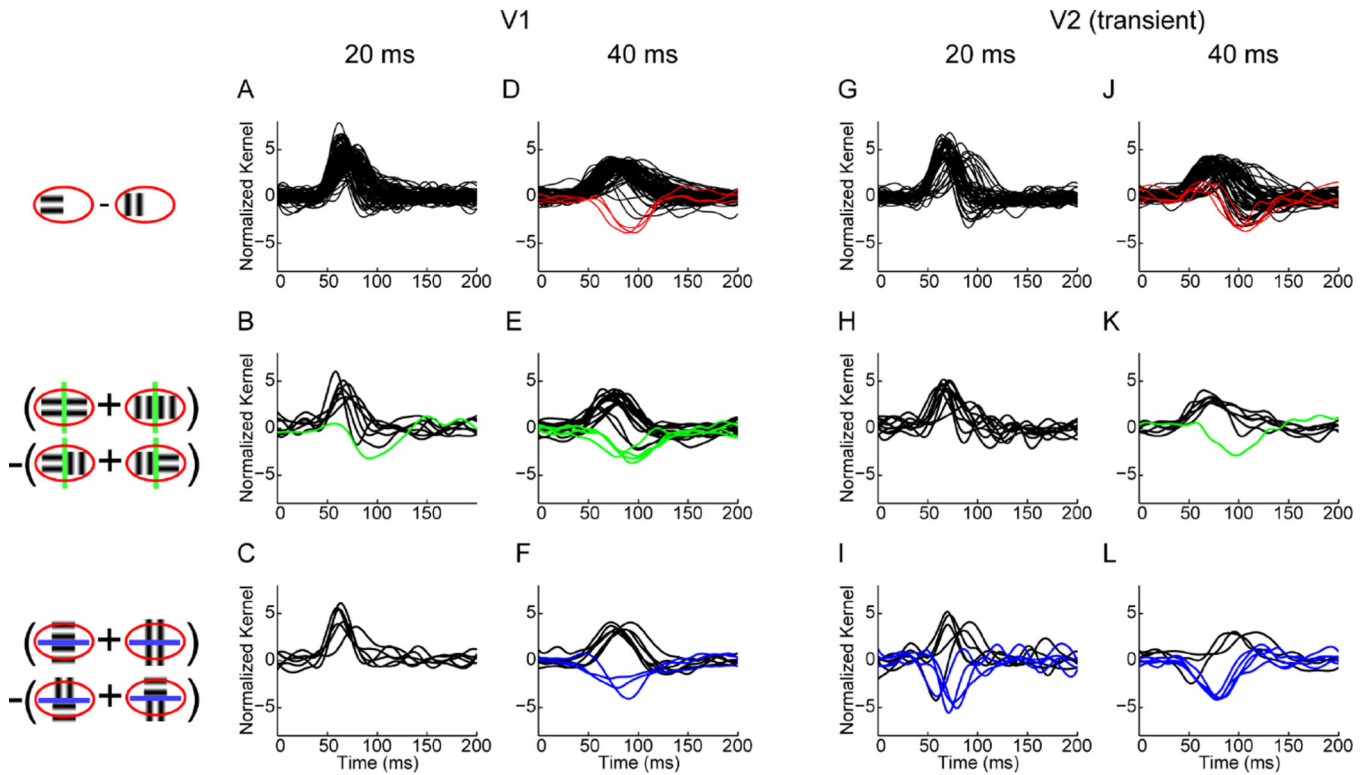


Figure 7. Responses of neurons to luminance and texture boundaries

Icons on the left illustrate computation of first- and second-order kernels (see also Figure 6B). All responses are normalized and positive responses are plotted in black. Negative responses are plotted in red for first-order (luminance boundary) responses, in green for second-order (texture boundary) responses orthogonal to the receptive field and in blue for those parallel to the receptive field. (A) First-order kernels (B) second-order kernels orthogonal and (C) second-order kernels parallel to the receptive field for V1 neurons at 20 ms frame duration. (D) First-order kernels (E) second-order kernels orthogonal and (F) second-order kernels parallel to the receptive field for V1 neurons at 40 ms frame duration. (G) First-order kernels (H) second-order kernels orthogonal and (I) second-order kernels parallel to the receptive field for transient V2 neurons at 20 ms frame duration. (J) First-order kernels (K) second-order kernels orthogonal and (L) second-order kernels parallel to the receptive field for transient V2 neurons at 40 ms frame duration. Adapted from (Schmid et al., 2014).