Response similarity as a basis for perceptual binding

Anna Sterkin
Goldschleger Eye Research Institute, Tel Aviv University, Tel Hashomer, Israel

Alexander Sterkin
Goldschleger Eye Research Institute, Tel Aviv University, Tel Hashomer, Israel

Uri Polat
Goldschleger Eye Research Institute, Tel Aviv University, Tel Hashomer, Israel

Detection of low-contrast Gabor patches (GPs) is improved when flanked by collinear GPs, whereas suppression is observed for high-contrast GPs. The facilitation resembles the principles of Gestalt theory of perceptual organization. We propose a model for contour integration in the context of noise that incorporates a temporal element into this spatial architecture. The basic principles are (1) the response increases with increasing contrast, whereas the latency decreases; (2) activity-dependent interactions: facilitation for low and suppression for high activity; (3) the variance increases with contrast for responses, rates, and latency; and (4) inhibition has a shorter time constant than excitation. When a texture of randomly oriented GPs is presented, the response to every element decreases due to fast inhibition between the neighboring elements, shifting the activity toward the range of collinear facilitation. Next, the slower excitation induces selective facilitation along the contour elements. Consequently, the response to the contour increases, whereas the variance of the rate and latency decreases, providing better temporal correlation between the contour elements. Thus, collinear facilitation increases the saliency of contours. Our model may suggest a solution to the binding problem by bridging between the temporal and spatial aspects of lateral interactions that determine the encoding of perceptual grouping.

Keywords: contour integration, grouping, binding, collinear facilitation, lateral interactions, temporal correlation, excitation, inhibition

Citation: Sterkin, A., Sterkin, A., & Polat, U. (2008). Response similarity as a basis for perceptual binding. *Journal of Vision, 8*(7):17, 1–12, http://journalofvision.org/8/7/17/, doi:10.1167/8.7.17.

Introduction

Grouping of small stimuli presented in the visual field into global objects is an important question that is under intensive investigation in contemporary neuroscience. Many studies have found principles that may reveal the roles of perceptual organization. These principles resemble the principles of the Gestalt theory of perception: the visual system groups features together based on similarity, proximity, smoothness, closure, common region, and connectedness (Koffka, 1935; for a review, see Rock & Palmer, 1990; Spillmann, 1997). According to these principles, the organization and the relationships between the local retinal inputs are the parameters that may determine the observer's performance in grouping global images.

Contour integration is a case of a grouping task. Many studies have attempted to understand the principles in which local features of the elements comprising the contour influence its detection. When the orientation between the two neighboring elements changes smoothly, the perception of continuity is produced. When several smoothly changed elements are arranged together, they may be perceived as a contour. However, similarity and smoothness alone may not be enough: similar elements arranged along the orthogonal axis to their local orientation may fail to form an easily detectable contour even if their local orientations are the same (Field, Hayes, & Hess, 1993; Polat & Sagi, 1993, 1994). It was found that the critical requirement for contour integration is the collinearity (smoothness, good continuation), i.e., the arrangement of the local elements along their local orientation (Kovács, 1996; Pettet, McKee, & Grzywacz, 1998). Other parameters may affect the detection of contours: the distance between the elements comprising the contour (proximity, closely arranged elements are better than far elements) and closed vs. open contour (Kovács, 1996; Kovács & Julesz, 1993). Taken together, integration of contours may be dependent on a combination of the local parameters of each element and the relationships between multiple elements. The weighting of each parameter may be not constant, but the overall combination of them may determine whether the contours are detected. Thus, sometimes one of the features may deviate from the basic rules, yet the combination of the other features may be strong enough for correct contour detection.

doi: 10.1167/8.7.17 Received October 15, 2007; published July 7, 2008 ISSN 1534-7362 © ARVO

It is still not clear how neurons in the visual system transform the principles of perceptual organization into neuronal code. It is well known that the response of individual neurons increases with increasing stimulus contrast (Albrecht, Geisler, Frazor, & Crane, 2002; Albrecht & Hamilton, 1982). The contrast response function (CRF) of neurons is generally described as a sigmoid non-linearity that accelerates for small contrasts and is saturated at high contrasts. It is also known that the response of neurons is maximal for the optimal orientation. However, many studies have shown that contour can be detected even when it is induced by collinearity of elements that are presented within noise of otherwise similar elements (Bonneh & Sagi, 1998; Field et al., 1993; Kovács, 2000; Kovács & Julesz, 1993; Pettet et al., 1998). Since the contrast of all the elements comprising both the contour and the background is the same, the output of the neurons responding to both contour and background should be the same. If so, the output of the individual neurons may not be sufficient to account for the observed grouping in contour integration.

Two general solutions to the binding problem were proposed (for a review, see Averbeck, Latham, & Pouget, 2006; Engel, Fries, & Singer, 2001). One is based on the spike rate of the neurons, assuming that neurons responding to contour should have a similar spike rate. As mentioned above, according to this approach, the spike rate of the neurons responding to the contour and background should be similar; thus, no differential information is carried in the spike rate per se. The second is by temporal correlation (synchronization) between responses of neurons responding to the contour. Although Usher and Donnely (1998) showed that binding is due to a global mechanism of grouping caused by temporal correlation of neural activation, the matter is still under debate (e.g., Dakin & Bex, 2002).

Accumulating evidence indicates that the contrast response is determined not only by the local stimulus parameters but is also modulated by the lateral interactions in the visual cortex. The contrast threshold and the response function of the neurons in the primary visual cortex are context dependent and can be modulated by remote image parts. The effects of context modulation, which may enable grouping of local elements into a global percept (Gestalt), was shown in psychophysical (Bonneh & Sagi, 1998; Herzog & Fahle, 2002; Kovács, 1996; Polat, 1999; Polat & Sagi, 1993, 1994) and physiological studies (Bauer & Heinze, 2002; Chavane et al., 2000; Kapadia, Ito, Gilbert, & Westheimer, 1995; Kapadia, Westheimer, & Gilbert, 2000; Levitt & Lund, 1997; Li & Gilbert, 2002; Mandon & Kreiter, 2005; Mizobe, Polat, Pettet, & Kasamatsu, 2001; Polat & Norcia, 1996; Schmidt, Goebel, Löwel, & Singer, 1997; Sengpiel, Baddeley, Freeman, Harrad, & Blakemore, 1998; Sillito, Grieve, Jones, Cudeiro, & Davis, 1995; Sugita, 1999, for a review, see Seriès, Lorenceau, & Frégnac, 2003). These studies clearly show that the neural representation of the target is modulated with regard to the surround stimuli. It is also apparent from these studies that the effect of center-surround is complex; it is mostly suppressive but may also be facilitative in some spatial-temporal combinations. The nature (either facilitation or suppression) and the strength of the context effect are determined by several parameters such as proximity, similarity, contrast, and global configuration.

Many algorithms have been proposed based on the available extensive physiological knowledge that incorporates the principle of integrating similar orientations along collinear directions, especially in the last decade (Field & Hayes, 2004). Every model incorporates a subset of features that characterize the visual processing in V1 and stimulus parameters that are of particular relevance for a given model. For instance, a lateral summation model was based on a network of lateral interactions that result in a deterministic output followed by a decision stage that utilizes probability summation over time and global threshold mechanisms to handle global noise (Usher, Bonneh, Sagi, & Herrmann, 1999). Other models also rely on the cumulative effect of local interactions over time (Li, 1998; Pettet et al., 1998) or on a combination of excitatory interactions with temporal synchronicity (Yen & Finkel, 1998). However, whereas Li (1998) has suggested how contour integration may be accomplished in V1 using the known elements of V1 with an optional add-on of a selective feedback mechanism from higher visual areas, Yen and Finkel (1998) conceived a cortical-based model for detecting contours embedded in noise. It is based on horizontal intra-cortical connections in V1-mediating context-dependent interactions and results in temporal synchronization. Another type of model postulated theorizes that contour integration may critically depend on the possible ability of the strength of local interactions to rapidly and selectively adapt to the local stimulus configuration (Braun, Niebur, Schuster, & Koch, 1994b). A new model of contour extraction in V1 was recently suggested (Ursino & La Cara, 2004). It differs from previous models since it incorporates four main mechanisms, according to recent physiological data: a feed-forward input from the LGN, an inhibitory feed-forward input, an excitatory cortical feedback, and a long-range feedback inhibition. The model extracts correct contours within acceptable time frames from image presentation (30–40 ms) and supports the notion that contour extraction is one the primary steps of visual processing, and that local processing in V1 is able to solve this task even under difficult conditions, without the participation of higher visual centers.

It was suggested earlier that there is an asymmetry between inhibition and excitation: whereas excitation develops slowly and is sustained, lagging behind the stimulus both at the onset and offset, inhibition is rapid and transient, following the onset and offset of the stimulus more precisely (Polat & Sagi, 2006; Polat, Sterkin, & Yehezkel, 2007). This suggestion is supported by the relatively slow time scale that characterizes lateral interactions (Bringuier, Chavane, Glaeser, & Frégnac, 1999; Grinvald, Lieke, Frostig, & Hildesheim, 1994; Seriès et al., 2003) and strong, transient, and fast inhibition (Bair, Cavanaugh, & Movshon, 2003; Borg-Graham, Monier, & Frégnac, 1998).

Based on the experimental findings described above, here we present a model emphasizing four fundamental features of the neuronal response:

1. the response increases with increasing contrast, whereas the latency decreases;
2. activity-dependent interactions: facilitation for low and suppression for high activity;
3. the variance increases with contrast for responses, rates, and latency; and
4. inhibition has a shorter time constant than excitation.

Methods

Stimuli

Neuronal responses to three types of stimuli are considered. The first is a single central Gabor patch (target GP), presented on a mean luminance background. The contrast is gradually increased from 1% to 100%; response amplitude and latency are the parameters of interest. In the second, two collinear GPs are added above and below the target GP (flanker GPs) with a spatial separation of 3 wavelenghts (λ), outside the classical receptive field of the neuron responding to the target GP (Polat, Mizobe, Pettet, Kasamatsu, & Norcia, 1998). The contrast of the flankers was constantly high (50%); otherwise, the flankers were identical to the target GP. The third is the case of contour integration: a texture of GPs, all the same contrast, equal spacing, random orientations (random noise), except for a path with the contour (see below).

Contrast response function

A typical neuronal contrast response function (CRF) was constructed using the Naka–Rushton equation that was earlier shown to provide the best fit to the measured responses of neurons in striate cortex of cats and monkeys (Albrecht & Hamilton, 1982):

$$R = R_{\max} \times \frac{c^n}{c^n + c_{50}^n}, \tag{1}$$

where R refers to response, $R_{\max}$ refers to the maximum response rate, c_{50} refers to the semisaturation contrast (the contrast required to produce 50% of the cell's maximum response), and n, the exponent, refers to the rate at which the changes occur. The parameters were set according to the measured values reported in the same study (Albrecht & Hamilton, 1982): $n = 3$, $c_{50} = 40$, $R_{\max} = 100$ (i.e., max response is 100%). For the condition of collinear modulation, the parameters were set according to the measured values reported in the cat's striate visual cortex (Polat et al., 1998): $n = 2$, $c_{50} = 25$, $R_{\max} = 70\%$ (due to inhibition for high contrast GPs by collinear flankers).

To calculate the variance of the response we used the equation:

$$V = b \times R^m, \tag{2}$$

where V refers to variance, R refers to response, m was set to 1, and b was set to 1.26 for a single GP condition and to 1.08 for the collinear modulation condition, according to the experimental data (Kasamatsu, Polat, Pettet, & Norcia, 2001; Polat et al., 1998). The decrease of variance following collinear modulation was included as an additional coefficient of a 2-fold decrease, compared with the variance of the response to the target GP for facilitation and a 2-fold increase for suppression (Kasamatsu et al., 2001; Polat et al., 1998).

$$V_{\text{collinear}} = \frac{1}{2} \times b \times R^m. \tag{3}$$

Contrast latency function

Typical neuronal contrast latency function was constructed using the inverted Naka–Rushton equation that was earlier shown to provide the best fit to the measured response latencies of neurons in striate cortex of cats and monkeys (Albrecht et al., 2002):

$$L = L_{\max} - L_{\text{shift}} \times \frac{c^n}{c^n + c_{50}^n} \tag{4}$$

where $L_{\max}$ is the maximum latency to the peak of the response at the lowest contrast, L_{shift} is the maximum possible shift (decrease) in latency, n is the latency shift exponent, c_{50} is the latency shift half-saturation, and c is contrast. It has been demonstrated that this function provides a good fit to the contrast-induced latency shifts (Albrecht, 1995; Albrecht et al., 2002). The parameters were set according to the measured values reported in (Albrecht et al., 2002): $n = 1.8$, $c_{50} = 24.6$ $L_{\max} = 121$, $L_{\text{shift}} = 65.3$. For the condition of collinear modulation, the parameters were set using a lookup table that directly related response to latency. The variance of the response latency was derived from the variance of the response rate, using the same lookup table.

Contour integration

For the contour integration case, we consider a texture of GPs of the same contrast, spatial frequency, and with equal spacing (3λ). All GPs are randomly orientated, except for a subset of elements whose orientation is arranged along a contour. As a result, a contour is embedded in a random noise.

It was shown earlier that maximal facilitation is found for collinear GPs (Polat & Sagi, 1994). Because facilitation is separation-dependent, 3λ was chosen as the spacing producing maximal facilitation (Polat & Sagi, 1993). The model computes the collinearity of each pair of GPs in the texture. Since the GPs are placed in an evenly-spaced matrix, each GP has 8 neighbors and 4 possible directions for collinearity (vertical, horizontal, and two oblique). The local orientation of each GP defines which direction of collinearity will be checked (GP is considered vertical, if its orientation deviates from the vertical direction by no more than 23 degrees; the rules for horizontal and oblique orientations are defined accordingly). For example, top and bottom neighbors of the middle GP in Figure 1A are checked for collinearity, since its local orientation is vertical (the middle GP and the two neighboring GPs are denoted by an ellipse). In another case, the local orientation of the middle GP is oblique (Figure 1B); therefore, only the two neighbors along the corresponding oblique direction are checked for collinearity (denoted by an ellipse). Thus, only directions in which global orientation matches the local orientation are considered, because only in this case collinearity is possible, and thus collinear facilitation may occur.

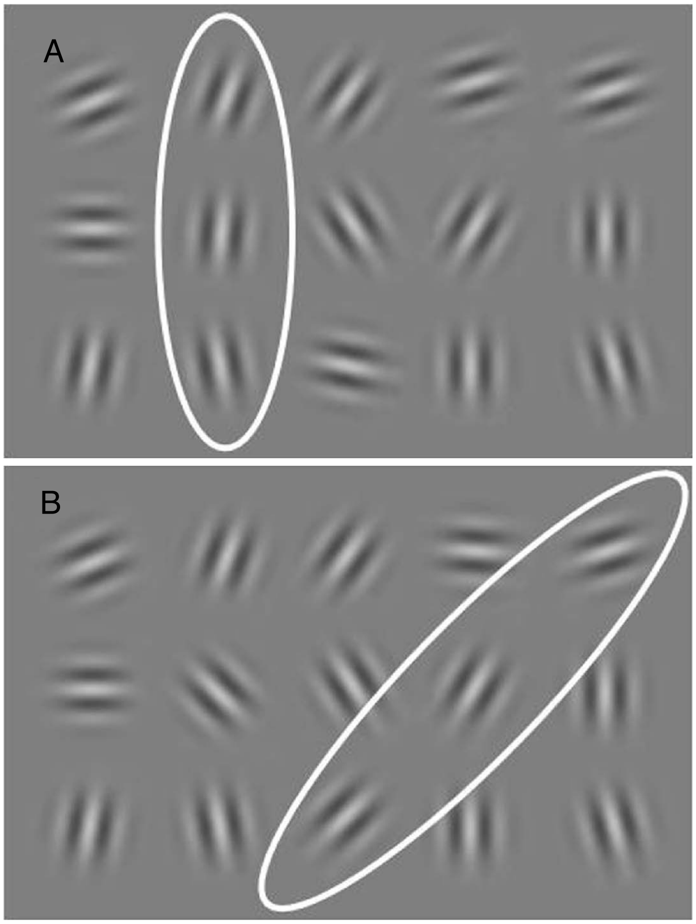

Figure 1. Matching global and local orientations: two examples. (A) A case of collinearity along the vertical direction. (B) A case of collinearity along the oblique direction.

Next, the rule for collinearity is applied: if the difference in orientation between the two analyzed neighboring GPs is less than 20 deg, they are considered collinear. Thus, the saliency of these two GPs is increased by 1 arbitrary level. This represents collinear facilitation of response.

This set of rules is applied in repeated steps in order to simulate consecutive iterations (or samplings) that are accumulated throughout the presentation of a visual stimulus. The orientation of the GPs that are part of the contour remains static, whereas the orientations of all other GPs randomly change in each iteration, in order to simulate the different noise patterns in each display of a real behavioral experiment. For instance, Kovács, Polat, Pennefather, Chandna, and Norcia (2000) used a different noise pattern of GPs in each of the tree displays, whereas the observers' task was to determine whether a closed contour was located on the left third, right third, or center third. It was also argued earlier that facilitation is sustained compared with fast inhibition (e.g., Polat & Sagi, 2006). It was shown that in monkey V1 the collinear facilitation associated with contour saliency developed much later than the background-induced inhibition (Li, Piech, & Gilbert, 2006). Therefore, iterative computation simulates the propagation of lateral interactions that induce collinear facilitation along the contour and contribute to its saliency in noise. The response of any two collinear GPs within the contour is increased with a 100% chance in each iteration, whereas the response of any two GPs in the background is increased with a much lower probability (by chance). Thus, with each iteration, the contour becomes increasingly salient. The facilitation gain in response to collinear GPs is accumulated upon consecutive iterations and eventually provides an adequate estimation of the contour path. The results of the iterations are summed up. We present a total of 30 iterations.

Iterations are also referred to in the literature as "time-step" (Braun, 1999) or "gist" (Joubert, Rousselet, Fize, & Fabre-Thorpe, 2007). It was argued earlier that presentation time is an important constraint for models in which contour salience results from an iterative computation, suggesting approximately 30 ms per time-step (Braun, Niebur, Schuster, & Koch, 1994a). Although this is far more that is needed for synaptic transmission, which is all that is postulated by some models of contour salience (Li, 1998; Pettet et al., 1998), it is just barely enough for establishing temporal synchronicity (Yen & Finkel, 1998), or for triggering NMDA-mediated plasticity (Braun et al., 1994a). The benefit of contour detection from longer durations of presentation was reported. Although detection was possible even for 80-ms durations (Polat & Bonneh, 2000), it deteriorated with the decrease of the presentation duration, which was directly compared by Field and colleagues (1993). In fact, it was much noisier for 250 than for 1000-ms durations. Moreover, a gist of a scene could be extracted with high accuracy (94% correct or more) from images that are flashed for about 20 ms by human observers (Delorme, Richard, & Fabre-Thorpe, 2000; Fabre-Thorpe, Delorme, Marlot, & Thorpe, 2001; Fabre-Thorpe, Richard, & Thorpe, 1998; Joubert et al., 2007; Rousselet, Macé, & Fabre-Thorpe, 2003; Thorpe, Fize, & Marlot, 1996; VanRullen & Thorpe, 2001). In our model, each iteration may represent a successive gist extraction, provided by longer presentation times that are typically allowed in the contour integration studies.

Divisive cross-orientation, contrast-dependent normalization of the response to a stimulus embedded in a noise, composed of randomly oriented stimuli, was included in the model as a scaling parameter *k*. *k* defines the normalization; *k* was set to 2 (for a detailed analysis of the quantitative dependence of normalization on contrast, see Heuer & Britten, 2002; for dependence of the suppression effect on contrast and *k,* see Chen, Kasamatsu, Polat, & Norcia, 2001; Kasamatsu et al., 2001; Sengpiel et al., 1998).

$$R_{\text{normalized}} = R_{\max} \times \frac{c^n}{k \times (c^n + c_{50}^n)}. \qquad (5)$$

Histograms

All histograms represent distributions of 100 iterations, divided in 20 bins for presentation. Each histogram was normalized to its total data count. Thus, the size of each bin represents a fraction of the data (in %, a total histogram sum equals 100). All histograms were fitted by a Gaussian function.

Results

It was previously shown that neuronal response is linearly related to the contrast of the stimulus for neurons in cat and monkey primary visual cortex (Albrecht & Hamilton, 1982); for a recent review, see (Albright & Stoner, 2002). The typical neuronal contrast response function (CRF) is best fitted by the Naka–Rushton equation (see Methods).

It was shown earlier that context may change the shape of the contrast response function depending on the spatial configuration of the stimulus (Chen et al., 2001; Kasamatsu et al., 2001; Mizobe et al., 2001; Polat et al., 1998; Sengpiel et al., 1998). Under conditions when the receptive field and the surrounds are stimulated by cross-orientation stimulus, the receptive field's response is suppressed by the surround in such a way that the CRF was shifted rightward (Sengpiel et al., 1998). This rightward shift is usually attributed to response normalization (Heeger, 1992; Sengpiel et al., 1998). However, under collinear interactions or iso-orientation configurations (Chen et al., 2001; Kasamatsu et al., 2001; Mizobe et al., 2001; Polat et al., 1998; Sengpiel et al., 1998), a change of the slope was observed. This change in the slope occurred as a result of increasing the contrast response near threshold (facilitation) and decreasing the response at higher contrast (suppression). It has been shown (Chen et al., 2001; Kasamatsu et al., 2001; Mizobe et al., 2001; Polat et al., 1998) that collinear interactions facilitate the response when the contrast is up to about twice the target's contrast threshold. The suppression dominated at contrasts that are above 8-fold the contrast threshold. Note that low and high contrast is a way to activate the neurons, and low and high contrasts are considered as low and high LGN activity (Polat, 1999; Stemmler, Usher, & Niebur, 1995).

Figure 2 shows the response of the neuron for a single central GP (red) and under collinear modulation (blue) as a function of increasing contrast of the central GP. Facilitation occurs for lower contrasts (near contrast

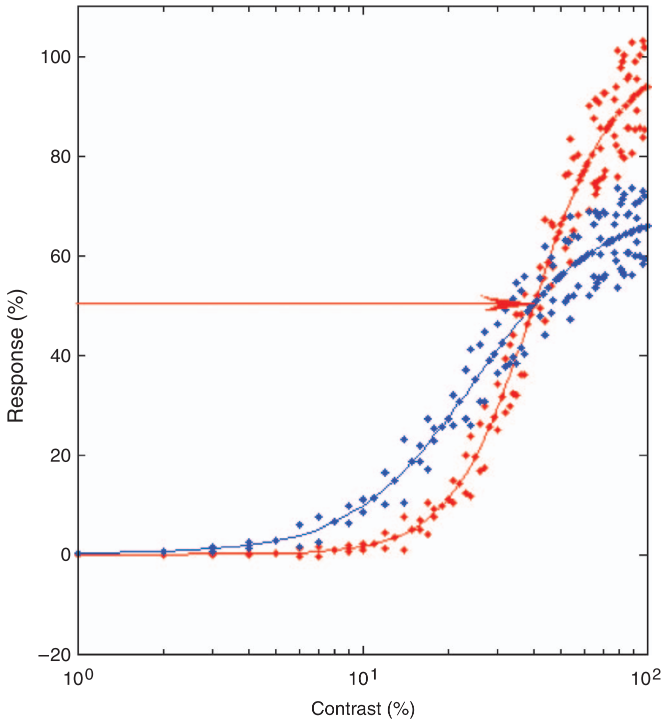

Figure 2. Contrast response function. Responses for a single central GP (red) and GP under collinear modulation (blue) as a function of contrast of the central GP are plotted on semi-log coordinates. The point of no modulation at the intercept of the two curves (cross-over, at 40% contrast) is denoted by a red arrow. The variance of response to a target GP is plotted with red dots and to target GP with collinear modulation with blue dots.

detection threshold) and suppression occurs at higher contrasts. The point of no modulation at the intercept of the two curves (cross-over) is observed at 40% contrast and is denoted by a red arrow.

The response of neurons in the retina and the lateral geniculate nucleus (LGN) is highly variable (noisy): repetitions of a constant stimulus may evoke a very different number of spikes. Thus, the input to the cortical neurons (EPSIs) is variable. A nearly constant ratio between the mean response and the mean variance was found under a variety of experimental conditions, suggesting that this is an innate property of these neurons (Geisler & Albrecht, 1995; Kasamatsu et al., 2001; Tolhurst, Movshon, & Dean, 1983).

Figure 2 shows that the response variance is linearly related to the response mean. The variance of response to a target GP (red dots) and to target GP with collinear modulation (blue dots) was calculated according to the experimental data (Kasamatsu et al., 2001; Polat et al., 1998). Figure 3 presents the response distribution for three contrasts—below the cross-over (10%), at the cross-over (40%), and above the cross-over (80%) for the two conditions—the single GP (red) and GP with collinear modulation (blue). As can be seen, the response is very variable even for the single GP. There is a large difference in the distribution of the responses around the mean observed for 100 iterations, which is lowest for 10% contrast and is dramatically increased toward 80% contrast. As a result of collinear modulation, the distribution is affected even though the mean response may be very similar (1.29 for the single GP vs. 1.39 for GP with flankers at 10% contrast, 45.07 vs. 45.73 at 40% contrast and 87.78 vs. 101.69 at 80% contrast), and there is a large difference in the distribution of the responses around the mean observed for 100 iterations. Although the distribution is much narrowed by collinear modulation at lower contrast (STD of 1.83 for the single GP vs. 0.93 for GP with flankers at 10% contrast), consistent with facilitation, it was not affected at the cross-over contrast (STD of 61.59 vs. 60.28 at 40% contrast), as predicted by no modulation, and it became wider at the contrast above the cross-over (111.74 vs. 194.65 at 80% contrast), consistent with suppression. Thus, the distribution of the mean is smaller in the collinear facilitatory interactions.

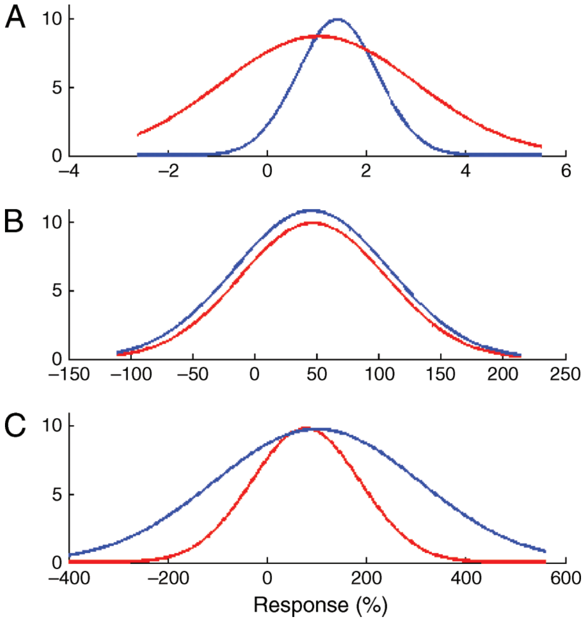

Figure 3. Collinear modulation of response. Histogram of response distribution for three contrasts: (A) below the cross-over (10%), (B) at the cross-over (40%), and (C) above the cross-over (80%) for the two conditions—the single GP (red) and GP with collinear modulation (blue).

An inverted relation between response latency and stimulus contrast was reported earlier for neurons in cat and monkey primary visual cortex (Albrecht et al., 2002) (Figure 3, red). Because there is large variance in the response amplitude to any given contrast, the latency is also variable, an effect that is derived from the distribution of the response (red dots).

To the best of our knowledge, there is no study that directly showed neuronal response latencies for a central GP with context modulation. Thus, here the latencies of the responses for a GP within collinear flankers are derived from the response function. A lookup table of the shift in latency that corresponds to the modulation of the response amplitude for each contrast was prepared. According to this table, latency as a function of contrast was plotted in Figure 4 (blue). Because the response for any given contrast is variable, there is a compatible distribution of the latencies for each contrast (blue dots). Moreover, we assume that, like in the case of the response amplitude, collinear facilitation decreases the variance, as compared to the distribution of response latencies for a single GP, by a factor of 2.

If we consider a contrast of 10%, which is below the cross-over and is facilitated by the collinear flankers, we can see that collinear modulation (blue) shortens the mean latency as compared to a single GP (red) (mean of 110.06 ms for the single GP vs. 95.71 ms for GP with flankers) but also dramatically narrows the distribution (STD of 4.84 ms for the single GP vs. 1.70 ms for GP with flankers) (Figure 5A). This is consistent with the experimental results that showed the reduction of variance following collinear facilitation (Kasamatsu et al., 2001). Based on the results above, we can assume that collinear modulation of the response latency may provide the grounds for synchronization between elements of a contour. For example, consider a case of two GPs in a contour that activate neurons with no overlapping of the receptive fields. The differences between the latencies (i.e., the delta latency) to the same contrast of 10% across multiple iterations will be considerably smaller for collinear elements than for non-collinear, as demonstrated in Figure 5B. Thus, we can conclude that the lower variance in the response latencies provides a much better temporal

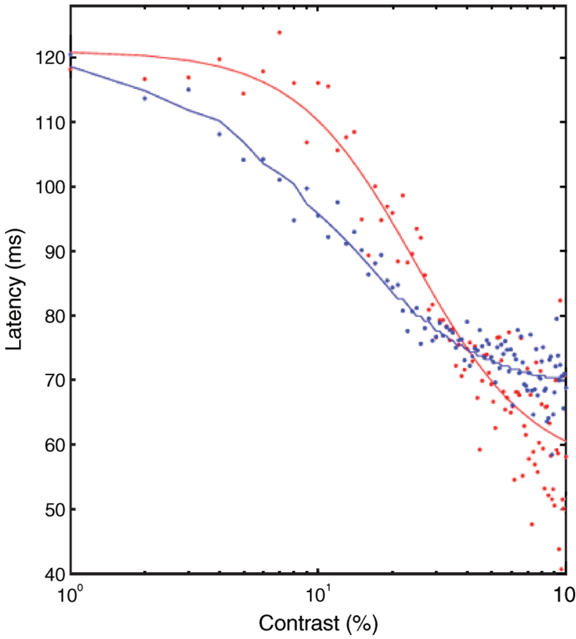

Figure 4. Contrast latency function. Response latency for a single central GP (red) and GP under collinear modulation (blue) as a function of contrast of the central GP are plotted on semi-log coordinates. The variance of response latency to a target GP is plotted with red dots and to target GP with collinear modulation with blue dots.

correlation between GPs that are subject to collinear modulation.

Our basic assumption is that inhibition is faster (i.e., has a shorter time constant) than excitation, based on real findings (Bair et al., 2003; Borg-Graham et al., 1998). Thus, when a texture of evenly spaced GPs of the same high contrast and random orientations is presented, the first stage in contour detection is inhibition. The CRF for every GP in the texture is shifted rightward (Figure 6A, green) whereas response distribution remains unchanged. Such suppression of responses within random noise was observed earlier and termed normalization (Heeger, 1992; Sengpiel et al., 1998). Evidence for normalization is seen in striate and extrastriate cortex from experiments where multiple stimuli are presented with a single receptive field. Neuronal responses in such experiments are smaller than that predicted by linear summation, revealing the presence of normalization.

It was recently shown that response is summed approximately linearly when contrast was low, but it rapidly becomes normalized as stimulus contrast is increased (Heuer & Britten, 2002). Moreover, responses recorded from monkey V1 (Li et al., 2006) support the notion that embedding a target in a complex environment leads to a shift toward a dominance of facilitation in collinear interactions, equivalent to the effect of a reduction in contrast (Kapadia, Westheimer, & Gilbert, 1999; Polat & Bonneh, 2000).

As was presented in Figure 2, the contrast at which the collinear modulation effect is reversed from facilitation to suppression is 40% (the cross-over, red arrow in both Figures 2 and 6A).

Figure 7A presents a demonstration of a contour consisting of local GPs (target) embedded in a similar background of GPs (noise). The only difference between GPs is their local orientation. When the orientation between the two neighboring elements changes smoothly (30 deg or less), it may enhance the perception of a contour. Several elements with smoothly changed orientation, arranged together, may form a contour. If we consider a texture of GPs at a constant contrast of 50%, we predict that an initial response that is evenly suppressed will be observed (Figure 6A). Therefore, for 3 sample contrast—below the crossover (30%), at the cross-over (40%), and above the cross-over (70%) (Figure 6B), when presented in a texture of randomly oriented GPs, the evoked response is reduced to the level that is below the cross-over, i.e., to the range of facilitation. Thus, in the second stage, the delayed facilitation increases the responses to the GPs that are defined as collinear by the model's assumptions, whereas the responses to the noise GPs remains suppressed.

We show that upon consecutive iterations the signal-to-noise ratio (SNR) is dramatically improved (Figure 7B). The response to the contour elements is considerably

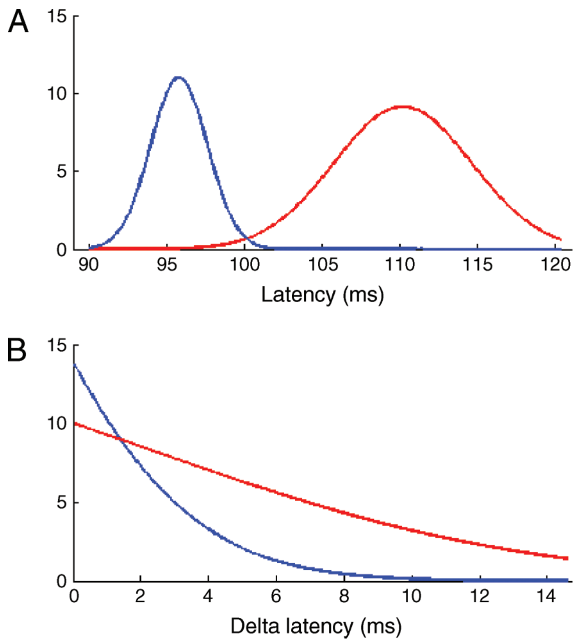

Figure 5. Collinear facilitation of response latency. An example of contrast of 10% (below the cross-over, i.e., in the range of facilitation) is presented. (A) Histogram of latency distribution is plotted for collinear modulation (blue) and a single GP (red). (B) Histogram of distribution of the delta latencies is plotted for a case of two GPs in a contour with either collinear (blue) or non-collinear (red) orientation.

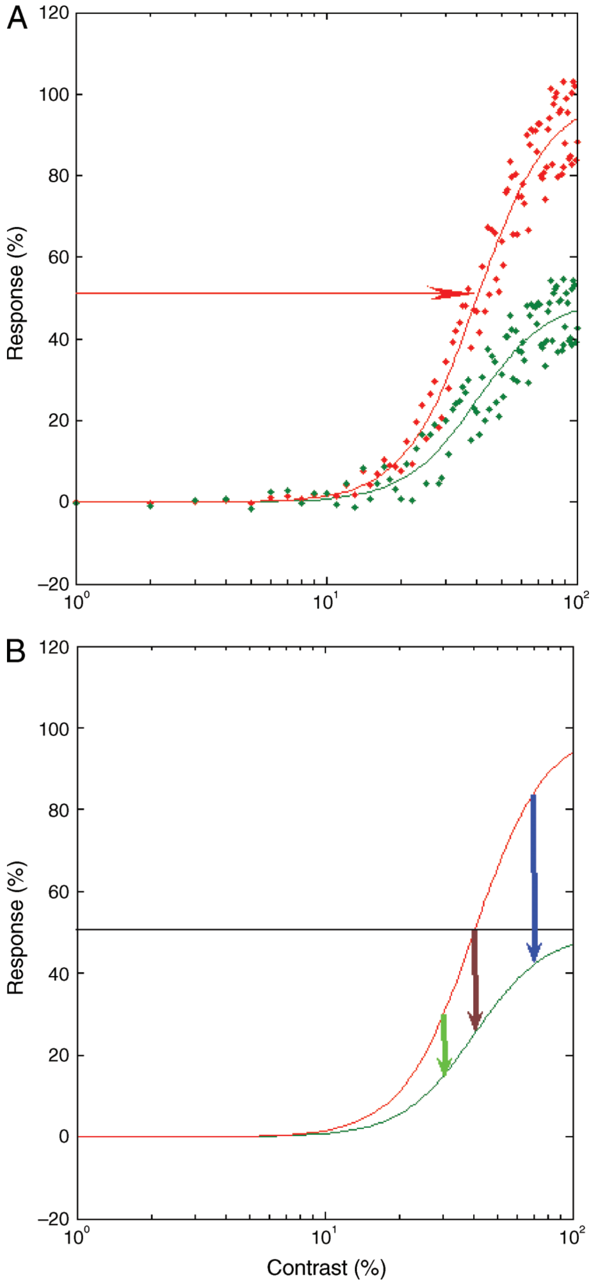

Figure 6. Contrast response function before and after normalization. (A) Responses for a single central GP (red) and GP after normalization by random noise (green) are plotted on semi-log coordinates. The point of collinear modulation cross-over (at 40% contrast) is denoted by a red arrow, as in Figure 1. The variance of response to a target GP is plotted with red dots and in green dots to target GP after normalization. (B) The same CRFs as in A are presented without the variance. The vertical arrows illustrate the effect of normalization for 3 sample contrast—below the crossover (30%, green arrow), at the cross-over (40%, brown arrow), and above the cross-over (70%, blue arrow).

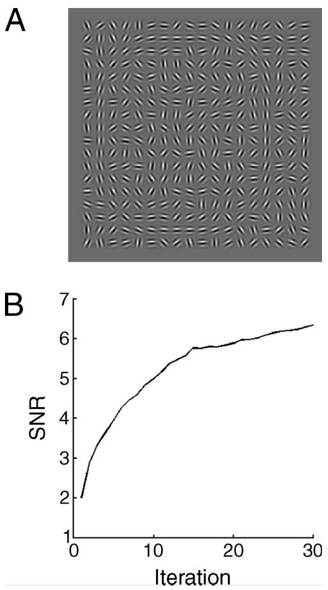

Figure 7. Contour integration in noise. (A) A demonstration of a contour consisting of local GPs (target) embedded in a similar background of GPs (noise). The path with the contour is defined by collinearity, i.e., by smooth change of orientation between each two neighboring elements (30 deg or less). (B) Signal-to-noise ratio (SNR) as a function of iteration.

higher than to noise and the contour is easily detected already by the third iteration, when compared to the 6th and 30th (Figures 8A–8D). In summary, neurons that respond to background elements will be suppressed by the network, whereas the response of the neurons responding to collinear elements of the contour will be amplified and eventually form the percept of a continuous contour.

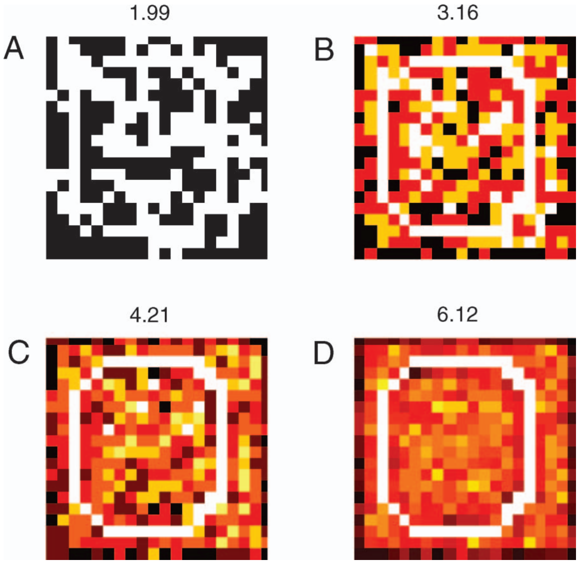

Figure 8. Signal-to-noise ratio (SNR) as a function of iteration. Response to a contour embedded in a noise background in the (A)1st , (B) 3rd , (C) 6th, and (D)30th iteration. SNR is specified for each panel.

Discussion

The main effect that has been shown in this study is that contour saliency is increased due to an increase of the response similarity between the contour elements in both the rate and temporal code. This effect is achieved because the response to each element in a contour is determined by the pooled facilitative and suppressive inputs from the elements comprising the smooth contour and the randomly oriented background elements, respectively. We showed that the response to each element alone is initially reduced (suppression) when it was embedded in a background of randomly oriented elements and that the response is then facilitated by collinear elements, even at the higher contrast levels. Thus, contour detection should be considered in terms of neuronal network but not on the basis of the output level of individual neurons responding to the discrete contour elements.

These results, in addition to those reported previously, clearly show that representation of contrast occurs in a spatially distributed fashion. Several computational models (e.g., Adini, Sagi, & Tsodyks, 1997; Stemmler et al., 1995; Yen & Finkel, 1998) have suggested a similar concept. Considering our network approach, it is more likely that the effect of collinear facilitation is to reduce the suppressive effect induced by the background. In other words, the delayed excitation between the collinear elements overcomes the fast inhibitory inputs from the background elements, thus increasing the contour saliency.

It is argued (Hess, Dakin, & Field, 1998) based on the results of Tolhurst (1989) that since the mean response and contrast are closely related, collinear facilitation at contrast of detection threshold cannot explain contour integration at supra-threshold contrasts. However, here we show that the output of an individual neuron is a conjunction of the response to the stimulus contrast and the modulatory effects of the lateral interactions. Thus, the neuronal output cannot be predicted solely from the target stimulus parameters. Our results here are consistent with the experimental data of Polat and Bonneh (2000).

Note that the perceived contrast of the target should not necessarily be changed due to the antagonistic effects of facilitation or suppression (Hess et al., 1998). It is possible that perceived contrast is based on the input level (first order) of the visual processing before the stage of contextual modulations (second order) where the contour integration may take place. Therefore, perception of global objects may be processed by the pattern of network activity of the neurons, each responding to the local feature of the image. The model predicts that this second-stage process may operate by correlating and matching the response activity of the neurons belonging to the same global pattern. Similarity in response is a new demonstration of the Gestalt similarity rule.

An important novel feature of this study is that the temporal response of the contour elements becomes less variable. As a result, the temporal correlation between the GPs within the contour elements is increased and reduced to within less than 3 ms (e.g., Gray, König, Engel, & Singer, 1989), a range that is found between synchronized neurons. Thus, the effect of collinear facilitation increases the similarity in both the rate and timing of responses between the contour elements. Support for this comes from the finding that the response of individual neurons in the cat primary visual cortex becomes more reliable (less variable and with a higher signal-to-noise ratio) due to collinear facilitation (Kasamatsu et al., 2001), and that phase lead has also been found in humans in conjunction with collinear facilitation (Polat & Norcia, 1996).

Although multiple parameters have been reported to affect contour integration in noise, such as uneven spacing between elements of contour and noise and contour closure (e.g., Braun, 1999; Kovács, 2000; Kovács & Julesz, 1993; Pettet et al., 1998; Yen & Finkel, 1998), in this study we considered only the elementary case of equal spacing between elements in the texture in order to reduce the number of parameters of the model and thus make a clear demonstration of the collinear facilitation effect on contour detection. However, the advantage of the current model is the combination of both the spatial and the temporal coding of contrast representation. It is probable that adding the other parameters with affect the performance of the model in a way similar to the effect on the performance of the observers in a real experiment.

Usually, studies of contour integration allow enough viewing time to achieve a satisfactory contour perception, using either static or a 1-sec-long presentation (Hess & Dakin, 1997; Kovács, 1996). Therefore, our model is realistic: each iteration represents an additional sampling of the visual stimulus within the allowed viewing time, thus allowing a reliable contour detection.

We have shown earlier that perceptual detection is completed in 100 ms and less (Polat et al., 2007). Moreover, it was shown that the critical duration for contour integration with collinear elements is performed in around 100 ms (Hess, Beaudot, & Mullen, 2001). The SNR in our model is sufficient for contour detection already by the third iteration which is within the usually used viewing time in contour integration experiments.

Conclusions

Taken together, the output of an individual neuron is not linearly related to the stimulus parameters and is highly susceptible to context modulations. Thus, even if the input contrast of an individual Gabor patch embedded in the contour is different from the others, the output of the neuron responding to it may be similar and become more reliable. Thus, the output of the neurons responding to the contour may share similar rates and temporal response properties and therefore may be more salient.

Theories on the neurophysiology of perceptual grouping have provided two alternative solutions to the binding problem: one is based on the spike rate of the neurons, assuming that neurons responding to contour should have a similar spike rate; the second is by temporal correlation (synchronization) between neurons responding to the contour. Our model may suggest a solution to the binding problem by bridging between the two theories, thus providing a link between the temporal and the spatial features of lateral interactions that determine the encoding of perceptual grouping.

Acknowledgments

We thank Prof. Misha Tsodyks for his perceptive comments on the revised version of the manuscript.

This research was supported by grants from the National Institute for Psychobiology in Israel, founded by the Charles E. Smith Family (UP and AS), the Israel Science Foundation, and the Nela Horovitz Foundation, TAU (UP).

Commercial relationships: none.

Corresponding author: Anna Sterkin.
Email: anna.sterkin@gmail.com.
Address: Goldschleger Eye Research Institute, Tel Aviv University, Tel Hashomer, Israel.

References

Adini, Y., Sagi, D., & Tsodyks, M. (1997). Excitatory-inhibitory network in the visual cortex: Psychophysical evidence. *Proceedings of the National Academy of Sciences of the United States of America, 94,* 10426–10431. [PubMed] [Article]

Albrecht, D. G. (1995). Visual cortex neurons in monkey and cat: Effect of contrast on the spatial and temporal phase transfer functions. *Visual Neuroscience, 12,* 1191–1210. [PubMed]

Albrecht, D. G., Geisler, W. S., Frazor, R. A., & Crane, A. M. (2002). Visual cortex neurons of monkeys and cats: Temporal dynamics of the contrast response function. *Journal of Neurophysiology, 88,* 888–913. [PubMed] [Article]

Albrecht, D. G., & Hamilton, D. B. (1982). Striate cortex of monkey and cat: Contrast response function. *Journal of Neurophysiology, 48,* 217–237. [PubMed]

Albright, T. D., & Stoner, G. R. (2002). Contextual influences on visual processing. *Annual Review of Neuroscience, 25,* 339–379. [PubMed]

Averbeck, B. B., Latham, P. E., & Pouget, A. (2006). Neural correlations, population coding and computation. *Nature Reviews, Neuroscience, 7,* 358–366. [PubMed]

Bair, W., Cavanaugh, J. R., & Movshon, J. A. (2003). Time course and time–distance relationships for surround suppression in macaque V1 neurons. *Journal of Neuroscience, 23,* 7690–7701. [PubMed] [Article]

Bauer, R., & Heinze, S. (2002). Contour integration in striate cortex. Classic cell responses or cooperative selection? *Experimental Brain Research, 147,* 145–152. [PubMed]

Bonneh, Y., & Sagi, D. (1998). Effects of spatial configuration on contrast detection. *Vision Research, 38,* 3541–3553. [PubMed]

Borg-Graham, L. J., Monier, C., & Frégnac, Y. (1998). Visual input evokes transient and strong shunting inhibition in visual cortical neurons. *Nature, 393,* 369–373. [PubMed]

Braun, J. (1999). On the detection of salient contours. *Spatial Vision, 12,* 211–225. [PubMed]

Braun, J., Niebur, E., Schuster, H. G., & Koch, C. (1994a). Perceptual contour completion: A model based on local anisotropic, fast-adapting interactions between oriented filters. *Society for Neuroscience Abstracts, 20,* 1665.

Braun, J., Niebur, E., Schuster, H. G., & Koch, C. (1994b). Perceptual contour completion: A model based on local, anisotropic, fast-adapting interactions between oriented filters. *Society for Neuroscience Abstracts, 20,* 1665.

Bringuier, V., Chavane, F., Glaeser, L., & Frégnac, Y. (1999). Horizontal propagation of visual activity in the synaptic integration field of area 17 neurons. *Science, 283,* 695–699. [PubMed]

Chavane, F., Monier, C., Bringuier, V., Baudot, P., Borg-Graham, L., Lorenceau, J., et al. (2000). The visual cortical association field: A Gestalt concept or a psychophysiological entity? *The Journal of Physiology, 94,* 333–342. [PubMed]

Chen, C. C., Kasamatsu, T., Polat, U., & Norcia, A. M. (2001). Contrast response characteristics of long-range lateral interactions in cat striate cortex. *Neuroreport, 12,* 655–661. [PubMed]

Dakin, S. C., & Bex, P. J. (2002). Role of synchrony in contour binding: Some transient doubts sustained. *Journal of the Optical Society of America A, Optics, Image Science, and Vision, 19,* 678–686. [PubMed]

Delorme, A., Richard, G., & Fabre-Thorpe, M. (2000). Ultra-rapid categorisation of natural scenes does not rely on colour cues: A study in monkeys and humans. *Vision Research, 40,* 2187–2200. [PubMed]

Engel, A. K., Fries, P., & Singer, W. (2001). Dynamic predictions: Oscillations and synchrony in top-down processing. *Nature Reviews, Neuroscience, 2,* 704–716. [PubMed]

Fabre-Thorpe, M., Delorme, A., Marlot, C., & Thorpe, S. (2001). A limit to the speed of processing in ultra-rapid visual categorization of novel natural scenes. *Journal of Cognitive Neuroscience, 13,* 171–180. [PubMed]

Fabre-Thorpe, M., Richard, G., & Thorpe, S. J. (1998). Rapid categorization of natural images by rhesus monkeys. *Neuroreport, 9,* 303–308. [PubMed]

Field, D. J., & Hayes, A. (2004). Contour integration and the lateral connections of V1 neurons. In L. M. Chalupa & J. S. Werner (Eds.), *The visual neurosciences* (vol. 2, pp. 1069–1079). MIT Press.

Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local "association field." *Vision Research, 33,* 173–193. [PubMed]

Geisler, W. S., & Albrecht, D. G. (1995). Bayesian analysis of identification performance in monkey visual cortex: Nonlinear mechanisms and stimulus certainty. *Vision Research, 35,* 2723–2730. [PubMed]

Gray, C. M., König, P., Engel, A. K., & Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature, 338,* 334–337. [PubMed]

Grinvald, A., Lieke, E. E., Frostig, R. D., & Hildesheim, R. (1994). Cortical point-spread function and long-range lateral interactions revealed by real-time optical imaging of macaque monkey primary visual cortex. *Journal of Neuroscience, 14,* 2545–2568. [PubMed] [Article]

Heeger, D. J. (1992). Normalization of cell responses in cat striate cortex. *Visual Neuroscience, 9,* 181–197. [PubMed]

Herzog, M. H., & Fahle, M. (2002). Effects of grouping in contextual modulation. *Nature, 415,* 433–436. [PubMed]

Hess, R. F., Beaudot, W. H., & Mullen, K. T. (2001). Dynamics of contour integration. *Vision Research, 41,* 1023–1037. [PubMed]

Hess, R. F., & Dakin, S. C. (1997). Absence of contour linking in peripheral vision. *Nature, 390,* 602–604. [PubMed]

Hess, R. F., Dakin, S. C., & Field, D. J. (1998). The role of "contrast enhancement" in the detection and appearance of visual contours. *Vision Research, 38,* 783–787. [PubMed]

Heuer, H. W., & Britten, K. H. (2002). Contrast dependence of response normalization in area MT of the rhesus macaque. *Journal of Neurophysiology, 88,* 3398–3408. [PubMed] [Article]

Joubert, O. R., Rousselet, G. A., Fize, D., & Fabre-Thorpe, M. (2007). Processing scene context: Fast categorization and object interference. *Vision Research, 47,* 3286–3297. [PubMed]

Kapadia, M. K., Ito, M., Gilbert, C. D., & Westheimer, G. (1995). Improvement in visual sensitivity by changes in local context: Parallel studies in human observers and in V1 of alert monkeys. *Neuron, 15,* 843–856. [PubMed]

Kapadia, M. K., Westheimer, G., & Gilbert, C. D. (1999). Dynamics of spatial summation in primary visual cortex of alert monkeys. *Proceedings of the National Academy of Sciences of the United States of America, 96,* 12073–12078. [PubMed] [Article]

Kapadia, M. K., Westheimer, G., & Gilbert, C. D. (2000). Spatial distribution of contextual interactions in primary visual cortex and in visual perception. *Journal of Neurophysiology, 84,* 2048–2062. [PubMed] [Article]

Kasamatsu, T., Polat, U., Pettet, M. W., & Norcia, A. M. (2001). Colinear facilitation promotes reliability of single-cell responses in cat striate cortex. *Experimental Brain Research, 138,* 163–172.

Koffka, K. (1935). *Principles of Gestalt psychology*. New York: Harcourt Brace.

Kovács, I. (1996). Gestalten of today: Early processing of visual contours and surfaces. *Behavioral Brain Research, 82,* 1–11. [PubMed]

Kovács, I. (2000). Human development of perceptual organization. *Vision Research, 40,* 1301–1310. [PubMed]

Kovács, I., & Julesz, B. (1993). A closed curve is much more than an incomplete one: Effect of closure in figure-ground segmentation. *Proceedings of the National Academy of Sciences of the United States of America, 90,* 7495–7497. [PubMed] [Article]

Kovács, I., Polat, U., Pennefather, P. M., Chandna, A., & Norcia, A. M. (2000). A new test of contour integration deficits in patients with a history of disrupted binocular experience during visual development. *Vision Research, 40,* 1775–1783. [PubMed]

Levitt, J. B., & Lund, J. S. (1997). Contrast dependence of contextual effects in primate visual cortex. *Nature, 387,* 73–76. [PubMed]

Li, W., & Gilbert, C. D. (2002). Global contour saliency and local colinear interactions. *Journal of Neurophysiology, 88,* 2846–2856. [PubMed] [Article]

Li, W., Piëch, V., & Gilbert, C. D. (2006). Contour saliency in primary visual cortex. *Neuron, 50,* 951–962. [PubMed] [Article]

Li, Z. (1998). A neural model of contour integration in the primary visual cortex. *Neural Computation, 10,* 903–940. [PubMed]

Mandon, S., & Kreiter, A. K. (2005). Rapid contour integration in macaque monkeys. *Vision Research, 45,* 291–300. [PubMed]

Mizobe, K., Polat, U., Pettet, M. W., & Kasamatsu, T. (2001). Facilitation and suppression of single striate-cell activity by spatially discrete pattern stimuli presented beyond the receptive field. *Visual Neuroscience, 18,* 377–391. [PubMed]

Pettet, M. W., McKee, S. P., & Grzywacz, N. M. (1998). Constraints on long range interactions mediating contour detection. *Vision Research, 38,* 865–879. [PubMed]

Polat, U. (1999). Functional architecture of long-range perceptual interactions. *Spatial Vision, 12,* 143–162. [PubMed]

Polat, U., & Bonneh, Y. (2000). Collinear interactions and contour integration. *Spatial Vision, 13,* 393–401. [PubMed]

Polat, U., Mizobe, K., Pettet, M. W., Kasamatsu, T., & Norcia, A. M. (1998). Collinear stimuli regulate visual responses depending on cell's contrast threshold. *Nature, 391,* 580–584. [PubMed]

Polat, U., & Norcia, A. M. (1996). Neurophysiological evidence for contrast dependent long-range facilitation and suppression in the human visual cortex. *Vision Research, 36,* 2099–2109. [PubMed]

Polat, U., & Sagi, D. (1993). Lateral interactions between spatial channels: Suppression and facilitation revealed by lateral masking experiments. *Vision Research, 33,* 993–999. [PubMed]

Polat, U., & Sagi, D. (1994). The architecture of perceptual spatial interactions. *Vision Research, 34,* 73–78. [PubMed]

Polat, U., & Sagi, D. (2006). Temporal asymmetry of collinear lateral interactions. *Vision Research, 46,* 953–960. [PubMed]

Polat, U., Sterkin, A., & Yehezkel, O. (2007). Spatio-temporal low-level neural networks account for visual masking. *Advances in Cognitive Psychology, 3,* 153–165.

Rock, I., & Palmer, S. (1990). The legacy of Gestalt psychology. *Scientific American, 263,* 84–90. [PubMed]

Rousselet, G. A., Macé, M. J., & Fabre-Thorpe, M. (2003). Is it an animal? Is it a human face? Fast processing in upright and inverted natural scenes. *Journal of Vision, 3*(6):5, 440–455, http://journalofvision.org/3/6/5/, doi:10.1167/3.6.5. [PubMed] [Article]

Schmidt, K. E., Goebel, R., Löwel, S., & Singer, W. (1997). The perceptual grouping criterion of colinearity is reflected by anisotropies of connections in the primary visual cortex. *European Journal of Neuroscience, 9,* 1083–1089. [PubMed]

Sengpiel, F., Baddeley, R. J., Freeman, T. C., Harrad, R., & Blakemore, C. (1998). Different mechanisms underlie three inhibitory phenomena in cat area 17. *Vision Research, 38,* 2067–2080. [PubMed]

Seriès, P., Lorenceau, J., & Frégnac, Y. (2003). The "silent" surround of V1 receptive fields: Theory and experiments. *The Journal of Physiology, 97,* 453–474. [PubMed]

Sillito, A. M., Grieve, K. L., Jones, H. E., Cudeiro, J., & Davis, J. (1995). Visual cortical mechanisms detecting focal orientation discontinuities. *Nature, 378,* 492–496. [PubMed]

Spillmann, L. (1997). Colour in a larger perspective: The rebirth of Gestalt psychology. *Perception, 26,* 1341–1352. [PubMed]

Stemmler, M., Usher, M., & Niebur, E. (1995). Lateral interactions in primary visual cortex: A model bridging physiology and psychophysics. *Science, 269,* 1877–1880. [PubMed]

Sugita, Y. (1999). Grouping of image fragments in primary visual cortex. *Nature, 401,* 269–272. [PubMed]

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381,* 520–522. [PubMed]

Tolhurst, D. J. (1989). The amount of information transmitted about contrast by neurones in the cat's visual cortex. *Visual Neuroscience, 2,* 409–413. [PubMed]

Tolhurst, D. J., Movshon, J. A., & Dean, A. F. (1983). The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Research, 23,* 775–785. [PubMed]

Ursino, M., & La Cara, G. E. (2004). A model of contextual interactions and contour detection in primary visual cortex. *Neural Networks, 17,* 719–735. [PubMed]

Usher, M., Bonneh, Y., Sagi, D., & Herrmann, M. (1999). Mechanisms for spatial integration in visual detection: A model based on lateral interactions. *Spatial Vision, 12,* 187–209. [PubMed]

Usher, M., & Donnelly, N. (1998). Visual synchrony affects binding and segmentation in perception. *Nature, 394,* 179–182. [PubMed]

VanRullen, R., & Thorpe, S. J. (2001). Is it a bird? Is it a plane? Ultra-rapid visual categorisation of natural and artifactual objects. *Perception, 30,* 655–668. [PubMed]

Yen, S. C., & Finkel, L. H. (1998). Extraction of perceptually salient contours by striate cortical networks. *Vision Research, 38,* 719–741. [PubMed]