



A unified account of tilt illusions, association fields, and contour detection based on elastica



Sander W. Keemink^{a,b,*}, Mark C.W. van Rossum^a

^aInstitute for Adaptive and Neural Computation, School of Informatics, University of Edinburgh, 10 Crichton Street, Edinburgh EH8 9AB, UK

^bBernstein Center Freiburg, Faculty of Biology, University of Freiburg, Hansastr. 9a, 79104 Freiburg, Germany

ARTICLE INFO

Article history:

Received 1 February 2015

Received in revised form 20 May 2015

Accepted 30 May 2015

Available online 22 August 2015

Keywords:

Association fields

Tilt illusion

Contextual interactions

Smoothness

Elastica

Gestalt

ABSTRACT

As expressed in the Gestalt law of good continuation, human perception tends to associate stimuli that form smooth continuations. Contextual modulation in primary visual cortex, in the form of association fields, is believed to play an important role in this process. Yet a unified and principled account of the good continuation law on the neural level is lacking. In this study we introduce a population model of primary visual cortex. Its contextual interactions depend on the elastica curvature energy of the smoothest contour connecting oriented bars. As expected, this model leads to association fields consistent with data. However, in addition the model displays tilt-illusions for stimulus configurations with grating and single bars that closely match psychophysics. Furthermore, the model explains not only pop-out of contours amid a variety of backgrounds, but also pop-out of single targets amid a uniform background. We thus propose that elastica is a unifying principle of the visual cortical network.

© 2015 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The Gestalt psychologists emphasized that human perception should be understood as a whole, rather than as the sum of individual elements. In the context of contour recognition, they proposed the law of good continuation, in which collinear or curvilinear line elements are associated together (Koffka, 1935; Wertheimer, 1923). This principle presumably underlies the ease with which humans extract smooth contours in natural as well as artificial images. The detection of contours does not necessarily require high level vision or receptive fields that encompass the complete contour. Instead, Field, Hayes, and Hess (1993) argued that association fields in early vision boost the response to collinear line elements, Fig. 1A, and that these local interactions are sufficient for contour detection. Computational models have used association fields to explain contour extraction and completion (e.g. Li, 1998, 1999; Tang, Sang, & Zhang, 2007; Williams & Thornber, 2001).

Evidence for the association field has been found in psychophysics (Field et al., 1993; Kapadia, Ito, Gilbert, & Westheimer, 1995; Kapadia, Westheimer, & Gilbert, 2000; Ernst et al., 2012), neural recordings (Bauer & Heinze, 2002; Kapadia et al., 1995, 2000; Li, Piëch, & Gilbert, 2006), and in neural

connectivity (Bosking, Zhang, Schofield, & Fitzpatrick, 1997), and supports the view that contour detection happens as early as primary visual cortex (V1). Furthermore, association fields have been linked to image statistics (e.g. Geisler, Perry, Super, & Gallogly, 2001; Sigman, Cecchi, Gilbert, & Magnasco, 2001) and predictive coding (Spratling, 2012).

From a functional point of view, association fields emphasize smooth contours. Indeed, in Fig. 1B the red bars are smooth continuations from the black bar, while the blue bars are more tortuous continuations. Smoothness can be quantified using the elastica principle, which measures the bending energy needed to connect two bars with a curve. In computer vision the curve with the lowest energy is better known as a spline (historically, a thin wooden rod used by draftsmen to create smooth curves). Line elements that can be connected with a low energy spline likely belong to the same object contour and are therefore particularly relevant to higher level vision. The elastica curve can also be seen as the maximum likelihood path of a stochastic contour completion process based on drifting particles, with the two bars as source and sink elements (Mumford, 1994; Williams & Jacobs, 1997; Williams & Thornber, 2001).

In this paper we propose that smoothness, as formalized by elastica, is the underlying principle for contextual modulations in V1. We first derive an efficient calculation of the elastica energy between two oriented line elements. We next assume that the contextual modulation in a neural population is determined by this

* Corresponding author at: Institute for Adaptive and Neural Computation, School of Informatics, University of Edinburgh, 10 Crichton Street, Edinburgh EH8 9AB, UK.
E-mail address: s.keemink@sms.ed.ac.uk (S.W. Keemink).

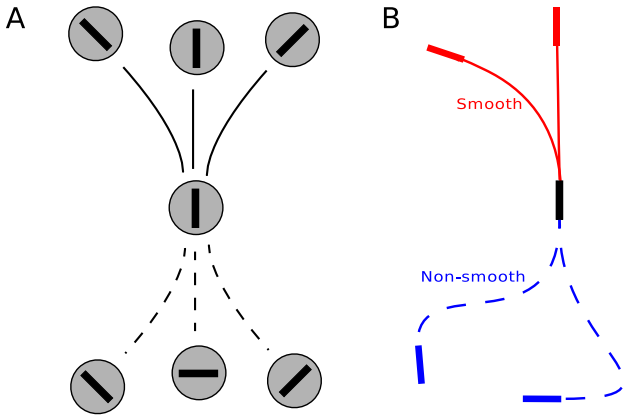


Fig. 1. Visual association fields. (A) The basic idea of an association field, as theorized based on psychophysical experiments in Field et al. (1993). Solid lines connect collinear elements and correspond to strong associations, while dotted lines correspond to weak associations. (B) Association fields and the relation to smooth contours. When trying to connect the center bar to any of the 4 presented flankers using smooth lines, the connections to the red bars are smoother and the connecting curve will have lower curvature energy than for the blue bars.

elastica energy. We find that the resulting model has (1) realistic association fields and neural responses, (2) produces various tilt illusions for both small sets of bars and grating-like stimuli, and, (3) leads to robust contour extraction, as well as single target pop-out from uniform backgrounds.

While links between these phenomena and elastica have been shown before individually, from contour extraction (Ernst et al., 2012; Sharon, Brandt, & Basri, 1997) to the tilt illusion (Schwartz, Sejnowski, & Dayan, 2006), there has been considerably variation in the precise implementation and biological realism of these studies. Here we present a unified account that, to our knowledge, is the first to collect such a wide range of phenomena in a single model with elastica at its core. Yet, the model is straightforward and biophysically realistic, relying only on independent contextual modulation terms. These results suggest that elastica is a core principle underlying the contextual interactions in V1.

2. Material and methods

2.1. Elastica

In this section we quantify smoothness according to the elastica principle. Extending earlier results we derive an accurate approximation for the smoothness of a curve connecting two line elements. Consider a scene with two bars, Fig. 2A: a center bar (red) and a flanker bar (blue). We define the positions and orientations as follows: The orientation of the center bar relative to the vertical is θ_c . The flanker is a distance r_f away from the center, and placed at a position which has an angle φ_f with the vertical. The orientation of the flanker is given by θ_f . The angles of the center and flanker bar relative to the line connecting them, are

$$\begin{aligned}\beta_c &= \varphi_f - \theta_c, \\ \beta_f &= \theta_f - \varphi_f,\end{aligned}$$

where the minus signs signifies circular differences, so that the angles lie in the interval $[-\pi, \pi]$.

We would like to know the smoothest curve connecting the two line elements. This minimization problem, known as elastica, has a long history dating back several hundred years (Mumford, 1994; Levien, 2008). In particular in computer vision and computer graphics this problem has been studied extensively, where the

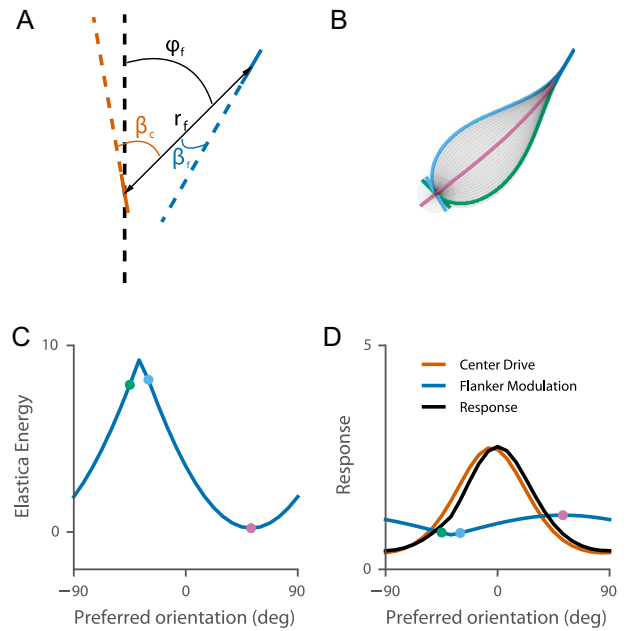


Fig. 2. The elastica smoothness and its proposed effect on surround modulation. (A) Setup of the center and flanker bars, illustrating the various angles used in the calculation of the curvature energy. (B) The family of minimum energy curves connecting the flanker to the center bar for various center orientations. (C) The energy of the curves in panel B, as a function of center orientation θ_c . The colored dots correspond to the colored curves in panel B. The cusp occurs when the minimum energy curve switches sides (between the light blue and green curve). (D) The population response of neurons with the center bar in A as their receptive field as a function of the neurons' preferred orientation. Each neuron receives a feed-forward drive from the center bar with a strength dependent on the neuron's preferred orientation (red). Each neuron also receives contextual modulation from the flanker (blue) which depends on the curvature energy of the elastica curve connecting the flanker to a center bar with the neuron's preferred orientation. The input drive is multiplied by the modulation to give the resulting population response (black), which is both deformed and shifted compared to the feed-forward drive.

smooth curves are known as splines. According to the elastica principle, the two line elements are imagined to be connected using a flexible rod. Since most human perception seems to be scale invariant, we use the scale invariant version of the elastica energy (Bruckstein & Netravali, 1990; Sharon et al., 1997)

$$E_{true} = L \int_0^L \left[\frac{d\Psi(s)}{ds} \right]^2 ds, \quad (1)$$

with s signifying the position along the curve, Ψ the relative angle of the curve at location s , and L the total length of the curve. The energy has a minimal value of zero when the curve is straight. As a curve becomes more tortuous, and thus has more curvature, the energy increases.

To find the smoothest curve, this energy has to be minimized w.r.t. $\Psi(s)$ and subject to the conditions $\Psi(0) = \beta_c$ and $\Psi(L) = \beta_f$. The length of the bars is assumed much shorter than r_f so that the curve between two bars goes from the middle of one bar to the other (see Section 4). The elastica energy can be found using the minimization method outlined in Sharon et al. (1997).

Typically elastica assumes a start and end direction. However, in our case, the elements have an orientation only, and we thus need a direction invariant energy, or the minimal energy across situations where the center and/or flanker angles can be flipped 180°. This results in the direction invariant elastica energy

$$E_{inv}(\theta_c, \theta_f, \varphi_f) = \min_{k,l \in \{-1,0,1\}} E(\theta_c + k\pi, \theta_f + l\pi, \varphi_f). \quad (2)$$

Fig. 2B shows the family of curves that minimize this energy as a function of the orientation of the center. The corresponding energy

for the curves is depicted in panel C. Notably, the energy has cusps where the solution switches from the green to the light blue curve.

Finding the true curvature energy is computationally expensive. However, the energy of the smoothest curve between a given center and flanker pair is approximated by [Leung and Malik \(2001\)](#) and [Sharon et al. \(1997\)](#)

$$E_s(\theta_c, \theta_f, \varphi_f) \approx 4(\beta_c^2 + \beta_f^2 - \beta_c \beta_f), \quad (3)$$

While Eq. (3) was derived from the assumption of small angles, it turned out to be a very good approximation for larger angles ([Sharon et al., 1997](#)). Furthermore, the direction invariant energy (Eq. (2)) based on this approximation is an exceptionally good match with the true direction invariant energy. [Fig. 3](#) compares the approximated and true energies across different center orientations for several flanker orientations. The error is small – certainly for our purposes – and free of qualitative differences.

2.2. Neural model

Here we implement how the neural responses in V1 are modulated by the surround using the elastica principle. For each bar, we assume that there is a population of $N = 32$ neurons, with preferred orientations ϕ_i in the interval $[-\frac{\pi}{2}, \frac{\pi}{2}]$. Initially we shall be mainly interested in the population that encodes the center bar. The response of a neuron in the absence of flankers is modeled by a von Mises function

$$g(\phi_i, \theta_c) = A_c \exp(K_c \cos 2[\phi_i - \theta_c]), \quad (4)$$

where A_c is the response amplitude, which we set to 1 Hz without loss of generality, as we only consider stimuli with identical contrast (see [Section 4](#) for a possible extension to stimuli with heterogeneous contrasts). θ_c is the orientation of the stimulus in the neuron's receptive field (the center bar). K_c sets the width of the neural tuning, with narrower curves for higher values. We use $K_c = 1$. The function $g(\cdot)$ across the population is illustrated in [Fig. 2D](#) by the red curve.

The smoothest curve connecting a flanker to a center bar with a neuron's preferred orientation ϕ_i has a curvature energy $E(\phi_i, \theta_f, \varphi_f)$. We propose that the neural response is modulated by a flanker through a modulation term

$$h(\phi_i, \theta_f, \varphi_f, r_f) = \exp \left[-\frac{a}{r_f} (E(\phi_i, \theta_f, \varphi_f) - E_0) \right].$$

This modulation is illustrated across the population (i.e. versus ϕ_i) in [Fig. 2D](#) by the blue curve. This formula was arrived at as follows. The elastica energy is always positive, but, in order to be consistent with physiology, we want flankers with a low curvature energy (i.e. high smoothness) to facilitate the response. Since $E \geq 0$ for all cases, we subtract an offset energy E_0 from the elastica energy, so that $h(\cdot) > 1$ for smooth contours. In addition, we divide the energy by r_f , the distance to the flanker, so that far away flankers do not modulate the response, similar to previous elastica studies ([Bruckstein & Netravali, 1990](#); [Sharon et al., 1997](#)). The gain parameter a determines the strength of the modulation. In the limit when $a = 0$, one has $h(\cdot) = 1$ and the response is independent of any flankers. For all simulations we use $a = 0.1$, and $E_0 = 4$. The exact values of these parameters have little qualitative effect, as will be discussed at the end of the results.

The modulation $h(\cdot)$ acts multiplicatively, consistent with physiology ([Cavanaugh, Bair, & Movshon, 2002](#)). We assume that each flanker contributes independently to the modulation, so that the final response of a neuron to a stimulus with n flankers is

$$r_i = g(\phi_i, \theta_c) \prod_{j=1}^n h(\phi_i, \theta_f^j, \varphi_f^j, r_f^j).$$

An example for $n = 1$ flanker is shown by the black line in [Fig. 2D](#).

2.3. Population vector

We read out the orientation encoded by the population using the population vector method. The population vector is a 2D vector given by the sum of the preferred orientation vectors of the neurons weighted by their firing rate ([Georgopoulos, Schwartz, & Kettner, 1986](#)), representing the estimated center orientation vector $\hat{\mathbf{v}}_c$

$$\hat{\mathbf{v}}_c = \sum_i r_i \mathbf{u}_i,$$

where r_i is the firing rate, and $\mathbf{u}_i = (\sin 2\phi_i, \cos 2\phi_i)$ is the unit vector pointing in neuron i 's preferred orientation (multiplied by two to ensure circularity). The estimated center orientation $\hat{\theta}_c$ follows from the angle of the population vector

$$\hat{\theta}_c = \frac{1}{2} \angle \hat{\mathbf{v}}_c, \quad (5)$$

where \angle denotes a vector's angle.

2.4. Computations

All data analysis and models were implemented in Python 2.7.5, using the Numpy 1.7.1, SciPy 0.12.0 and Matplotlib 1.2.1 toolboxes. Model code is available at [<https://github.com/swkeemink/elastica>].

3. Results

In the context of the Gestalt law of good continuation, we study a neural network in which the contextual interactions are based on pairwise optimal smoothness. For clarity we restrict ourselves to images composed of bars with various orientations all with identical contrast, and at a single spatial scale. Each bar is assumed to fall in the classical receptive field of a population of neurons with preferred orientations $\phi_i = -\pi/2 \dots \pi/2$. We will refer to the bar under consideration as the center bar, and the other bars as the flankers. As an example, consider the scene in [Fig. 2A](#), where the red bar represents the center and the blue bar a single flanker. We write the neural response of neuron i with preferred orientation ϕ_i as

$$r_i = g(\phi_i, \theta_c) h(\phi_i, \theta_f, \varphi_f, r_f).$$

The function $g(\cdot)$ models the classical V1 orientation response to the center stimulus using a von Mises function, where θ_c is the orientation of the center stimulus. The function $h(\cdot)$ models the contextual modulation from the flanker as follows ([Section 2](#)):

$$h(\phi_i, \theta_f, \varphi_f, r_f) = \exp \left\{ -\frac{a}{r_f} [E(\phi_i, \theta_f, \varphi_f) - E_0] \right\},$$

where θ_f, φ_f and r_f are the flanker orientation, angular position, and distance from the center respectively. $E(\cdot)$ describes the curvature energy of the smoothest curve connecting the flanker to a bar of orientation ϕ_i in the center. The parameters a and E_0 set the strength and offset of the modulation (see [Section 2](#) and below). The smoother the curve from the flanker to the center, the more positive the modulation (for smooth curves for which $E(\cdot) < E_0$ it become facilitatory, $h(\cdot) > 1$), whereas when the curve is tortuous and thus has a high elastica energy, the modulation is inhibitory. In [Section 2](#) the functional form of the modulation is derived from the elastica principle, and an efficient approximation for the curvature energy is presented. The center drive $g(\cdot)$, contextual modulation $h(\cdot)$ and resulting population response \mathbf{r} are illustrated in [Fig. 2D](#). It can be observed that due to the modulation the population response is a deformed version of the center drive.

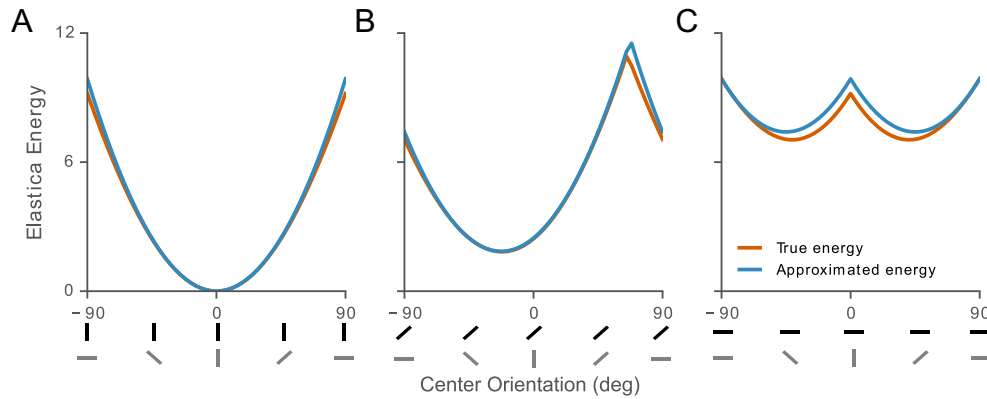


Fig. 3. The Sharon approximation (blue curve, Eq. (3)) compares well to the true elastica energy (red curve, Eq. (1)). Both energies were made direction invariant (Eq. (2)). (A–C) The elastica energy across different center orientations, for fixed flanker orientations of 0° , 45° and 90° respectively. For -90° , -45° , 0° , 45° and 90° a situation sketch is shown, with the gray bar indicating the center orientation, and the black bar the flanker orientation. Note that for a perpendicular flanker, relative to the center, the energy is minimal at two center orientations resulting in two dips in the energy, panel C.

3.1. Single flanker association field

We first analyze the association field that the model predicts by considering the effect of a single flanker on the response of a neuron with preferred orientation $\phi_i = 0$, as represented by the black bar in Fig. 4. A flanker in a given surround location will modulate the response of this neuron by a factor $h(\phi_i = 0, \theta_f, \varphi_f, r_f)$. In Fig. 4A we indicate at each location which flanker orientation yields the most positive modulation, $\text{argmax}_{\theta_f} h(\phi_i = 0, \theta_f, \varphi_f, r_f)$. The color of a bar indicates excitation (red, $h() > 1$) or inhibition (blue, $h() < 1$), and its opacity is proportional to the modulation strength. The bars that increase the response (red bars in panel A), correspond to flankers that form a smooth contour. Flankers with an inhibitory effect correspond to more tortuous contours. The parameter E_0 determines for which amount of curvature the modulation become inhibitory (see also Fig. 8).

To get further insight in the model we plot the bar which leads to the most suppressive modulation, $\text{argmin}_{\theta_f} h(\phi_i = 0, \theta_f, \varphi_f, r_f)$, which can be called the ‘dis-association field’ in panel B. Consistent with Fig. 3, the strongest suppressing flankers are approximately rotated 90° from the most facilitating ones. Furthermore, inhibition exists across a wider range of orientations.

Finally, we plot the effect of flankers oriented the same way as the preferred orientation to more directly test our association field against neural measurements, in which the flanker orientation is usually kept the same as the preferred orientation of the neuron being measured, panel C.

The butterfly shapes in panel A and C corresponds to both psychophysics in the form of the association field (Field et al., 1993), as well as electrophysiological results in monkeys where, in particular at low contrast, flankers collinear with the preferred orientation in the center excite (Kinoshita, Gilbert, & Das, 2009), and flankers parallel to the center inhibit (Kapadia et al., 1995, 2000).

3.2. The tilt illusion: two flankers

Next we study the case when two flankers are placed oppositely each other, around the center. To extend the model to this situation we assume that each flanker independently modulates the response, thus the response of a neuron in the center population is $r_i = g(\phi_i, \theta_c)h(\phi_i, \theta_f^1)h(\phi_i, \theta_f^2)$. We decode the center population response using the population vector (Section 2). Due to deformation of the population response caused by the modulation, the population vector is no longer aligned with the stimulus orientation. Psychophysically this presumably leads to a tilt illusion in the percept of the center bar.

We illustrate several configurations in Fig. 5. In each panel we show two example rotations in the top row. The elastica curves connect the flankers to curves in the center with different preferred orientations; the curve’s opacity is proportional to that neuron’s response, indicative of the resulting population response.

For parallel flankers rotating around the center, panel A, the elastica energy is smallest for neurons with preferred orientations $\pm 45^\circ$ from the flankers’ orientations Fig. 3C, resulting in two

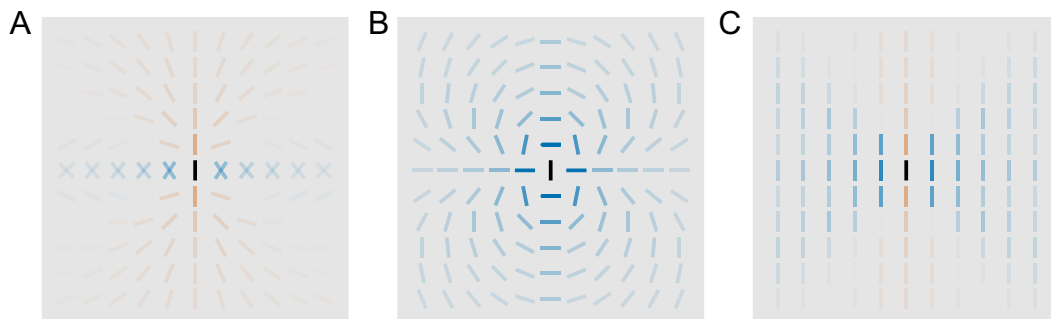


Fig. 4. Association fields as derived from the elastica principle. (A) The association field for a neuron with preferred orientation indicated by the black bar. In each location, the flanker orientation that most positively modulates the activity is shown. Red (blue) flankers correspond to facilitatory (inhibitory) modulation; the opacity indicates the strength of the modulation. (B) Same as in A, but instead the orientations that most **decrease** the activity are shown (the “dis-association field”). (C) The effect of a flanker of the same orientation as the preferred orientation in the center.

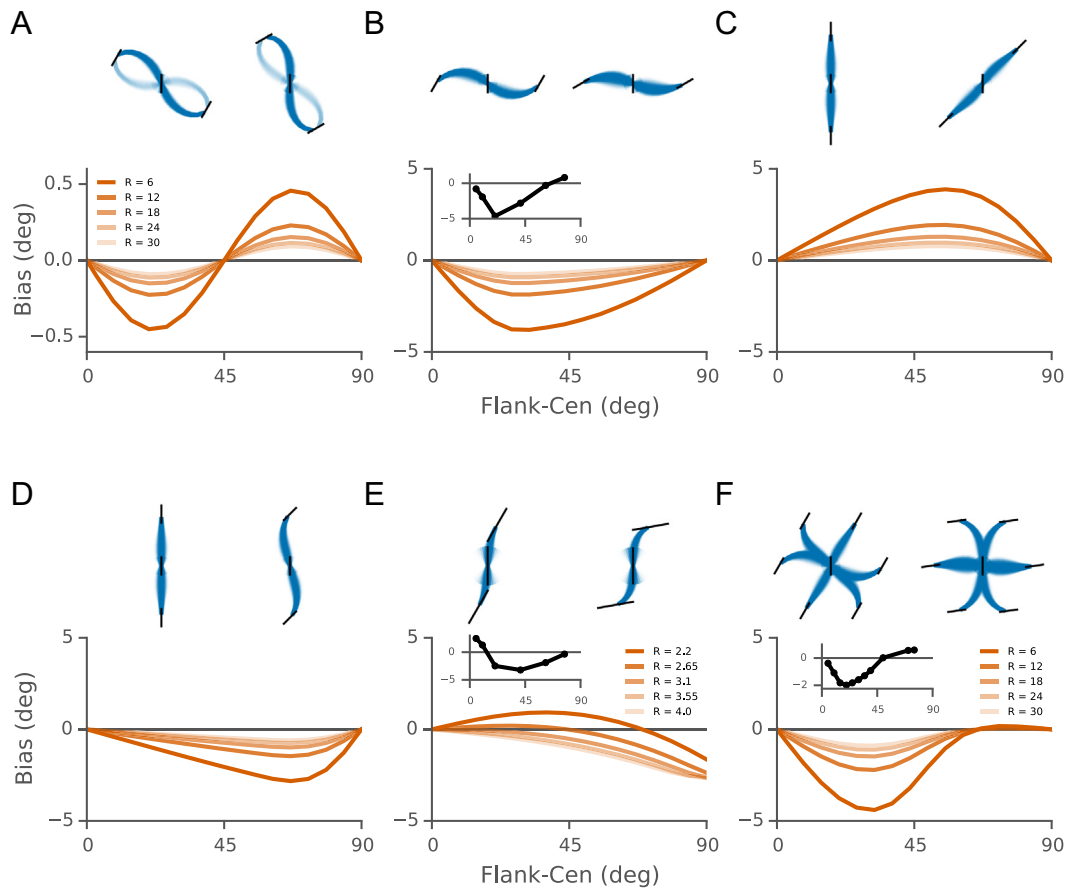


Fig. 5. Tilt illusions from elastica-based contextual modulation by flanker pairs. Negative bias indicates repulsion, positive bias indicates attraction. Top rows: example configurations, the elastica curves connect the flankers to neurons in the center population with a specific preferred orientation. The opacity is proportional to the response of the corresponding neuron. (For clarity the modulation strength was increased to $a = 0.5$.) Bottom rows: tilt bias as a function of flanker orientation. (A) Two parallel lateral flankers are rotated around the center (at 0° the flankers are on the sides). The illustrated flanker tilts are 30° and 60° from left to right. (B) Same as panel A, but with the flankers kept in place. The tilts are 30° and 60° left and right. Inset: illusion data extracted from Kapadia et al. (2000), Fig. 4C. (C) Two aligned flankers, rotating around the center. The examples have a rotation of 0° and 45° respectively. The elastica curves follow the flanker tilt and then bend back towards the center, resulting in an attractive illusion. (D) Keeping the flankers in place and tilting them in place. The elastica curves follow the flanker tilt and then bend back towards the center. (E) Reproducing the alignment in experiment in Kapadia et al. (2000), the ends of the flankers were aligned with the center bar. Inset: illusion data extracted from Kapadia et al. (2000), Fig. 4C. (F) Reproducing the setup in Westheimer (1990), with 6 flankers in a hexagon which are rotated in place. Inset: averaged illusion data extracted from Westheimer's Fig. 1.

“valleys” in the energy. Accordingly, the elastica curves from these flankers fan in two directions, Fig. 5A top. The resulting modulation shifts the population response depending on the flanker orientation. In the example of a 30° flanker rotation, top left, the modulation is mostly counter-clockwise. However, when the flanker is rotated to 60° , top right, the same curves pass the vertical before reaching the center, resulting in the population responses being shifted clockwise. These effects result in a repulsive illusion for flanker rotations 0 – 45° and an attractive tilt for 45 – 90° , panel A bottom.

When the lateral flankers are rotated in place, Fig. 5B top, the effect is always repulsive. With the flankers either at 30° (left) or 60° (right), the most influential elastica curves from the left flanker always move up first before going down (and vice versa for the right flanker), ending in an orientation which is repulsed away from the flanker orientation.

For the aligned flankers rotating around the center, panel C top, elastica curves connect smoothly to the flankers, with the lowest energy curve being a straight line. As a result the modulation is excitatory for neurons with a preferred orientation close to the flanker orientation (Fig. 4A), and suppressive far away, and the population response follows the flankers. Across flanker orientations this generates an attractive illusion. Note that the attractive effect does not require excitation, it is sufficient that

the modulation from both flankers is least suppressive for the neurons with a preferred orientation equal to the flanker orientation.

Next, we rotate the flankers in place above and below the center, panel D top. When the flankers are tilted 45° clockwise, right example, the elastica curves from the top flanker start off towards the left, then bend back towards to center. The population response is thus shifted away from the flanker orientation, resulting in a repulsive illusion.

The two bar illusions in the model are in close accordance with known psychophysics. Westheimer (1990) reported repulsion from tilted flankers on the sides (Fig. 5B), or above and below (panel D). The orientation dependence of the repulsion from lateral flankers (panel B) closely matches (Kapadia et al., 2000) who also found repulsion maximal at around 30° (except for the small attractive effect they reported for larger flanker orientations), inset panel B. Kapadia et al. (2000) also placed flankers above and below the center, and tilted them (as in panel D). But in addition the flankers were displaced so that the near end of the flankers aligned with the center bar. In this case they found an attractive effect for small flanker orientations, and a repulsive effect other orientations, inset panel E. The attractive effect went away quickly as the flanker distance was increased.

We reproduced their setup by assuming a finite bar length for the flanker rotation, and displacing the flanker accordingly, as in

Fig. 5E top. Since the bar lengths in our model are assumed infinitesimally short, the curves are drawn from and to the centers of the bars. This turns out to be key to understanding the attractive part of this illusion. As long as the flankers are close enough, smaller orientations will seem aligned as the origin of the elastica curves is slightly displaced, panel E top left. When the flanker orientation is larger, this is no longer the case and the curves first need to pass the vertical before reaching the center, resulting in a repulsive effect. When the flankers are moved further away (i.e. displaced in the y direction), the situation becomes more analogous to panel D, as the x -displacement of the flanker becomes insignificant, and the illusion turns repulsive for all flanker orientations. Although the precise angular dependence in the model does not match the data, it is surprising that it can exhibit both tilt effects.

All illusions and modulations weaken as the flanker distance was increased, as illustrated by curves of decreasing opacity in the bias graphs. This is in accordance with most tilt illusion studies which note that the illusions decreases in strength as the contextual stimuli are placed further way (Kapadia et al., 2000; Westheimer, 1990).

In the context of elastica, the illusions in Fig. 5B and D were previously explained by noting that the lowest energy curve connecting the two flankers is either oriented towards the flankers' orientation, as in A, or away from them, as in B (Schwartz et al., 2006). Bayesian estimation then results in an orientation between this smooth orientation and the presented center orientation, resulting in the repulsive or attractive illusion. However, an additional displacement of the center bar was allowed to produce an attractive solution for small angles in panel C.

3.3. The tilt illusion: full surround

The two flanker stimuli described above, lead to both repulsive and attractive effects. However when either a hexagon of surround bars (Westheimer, 1990), or a surround grating (Clifford, 2014) is used a repulsive tilt illusion occurs for small center-surround orientation differences, while a weak attractive effect occurs for larger orientation differences (inset Fig. 5F), which has been speculated to have a different origin (Clifford, 2014). Mechanistically, the repulsive tilt illusion has been explained by the fact that a surround grating results in orientation tuned suppression, with most suppression when the surround is the same as a neuron's preferred orientation (e.g. Clifford, Wenderoth, & Spehar, 2000; Schwartz, Hsu, & Dayan, 2007).

To examine these illusions in our model, we first turn to Westheimer's experiment, which we can reproduce exactly. There a hexagon of 6 flankers was placed around a center bar, evenly spaced so that there are two parallel bars on the sides, as in Fig. 5F top. The bars were then rotated in place and we measured the effect on the neural population response at the center location. As above we assume that each flanker independently modulates the center responses and the orientation of the center bar was again decoded from the neural activities using a population vector. For most orientations, the net effect from the elastica curves to the center is repulsive. However, when they are close to perpendicular to the center, the four top and bottom flankers win out with a small attractive effect (also see Fig. 8). Thus the model explains both the repulsive illusion, and the attractive effect for larger orientation differences.

We next approximate a center grating by a single oriented bar, and a surround grating as a large set of 16 identically oriented bars, Fig. 6A. At first glance perhaps a weak attractive tilt would be expected again. However, the net modulation from all flankers is inhibitory, panel B, and strongest when a neuron's preferred orientation is the same as the surround orientation, in close

accordance with known neural responses. As a result, the decoded orientation is repulsed away from the surround orientation, corresponding to a repulsive tilt illusion, panel C. We also varied the number of flankers and found much the same effects, with still a weak attractive effect for 8 bars, but repulsion otherwise (not shown). In summary, tilt illusions in stimuli with surround gratings and with pairs of flankers can be unified under the elastica principle.

3.4. Contour detection

So far we have focused on the effect of flankers on the decoded orientation of a center bar. We now turn our attention to larger scenes consisting of several bars, where we find the response to each bar in succession by taking that bar as the center, and considering all other bars as the flankers. The principle of elastica and smooth contours has classically been used to extract contours from images. These implementations typically explicitly calculate the curvature energy for all element combinations, rather than incorporating the energy in a modulation term as we do here. While association fields in general have been used to facilitate contour detection through contextual modulations (e.g. Bauer & Heinze, 2002; Field et al., 1993; Li, 1998, 1999), here we examine if elastica based modulation also leads to contour extraction.

To study contour extraction in our model we measure the apparent saliency s of the contour (as in Li, 1999). The apparent saliency of a bar is defined as the maximal response in the population responding to that bar (i.e. the maximum over their preferred orientations ϕ_i), compared to the maximal responses to bars in other locations

$$s_b = \frac{\max_{\phi_i}(r)}{\langle \max_{\phi_i}(r) \rangle_{\text{image}}},$$

where the average in the denominator is taken over the whole image. The saliency of a complete contour is defined as the mean maximal response to a contour relative to the whole image

$$s = \frac{\langle \max_{\phi_i}(r) \rangle_{\text{contour}}}{\langle \max_{\phi_i}(r) \rangle_{\text{image}}},$$

where the average in the nominator is over the bars that constitute the contour. If $s > 1$, the responses are higher in the contour and it is salient. We also implemented a mean based saliency measure, which uses mean responses instead of maximal responses. This resulted in weaker saliencies, but no qualitative differences (not shown).

In the plots that follow, we show the encoded image above its modeled percept. The opacity of the decoded bars is proportional to their saliency. Consider, first, a lone target bar amidst a homogeneous background, Fig. 7A. While this is not a true contour, from psychophysics and neural measurements we expect the target to be salient (Nothdurft, 1993; Shushruth et al., 2013). Indeed, the neurons in the background are inhibited more than the center flanker, similar to the effect in Fig. 6B. Due to the resulting higher neural response of the center bar it jumps out from the background, as signified by its darker color. As the target is rotated towards the surrounding orientations, the saliency decreases until it is no longer salient (not shown).

Next, we embed a simple contour in a homogeneous background, as in Fig. 7B. Here again, for the decoded image the bars of the feature of interest are darker than those in the background, indicating a salient contour. In this case, although actually all bars in the image experience suppression, those in the background are suppressed more, since they are surrounded by more bars of similar orientations. However, bars that are part of the contour

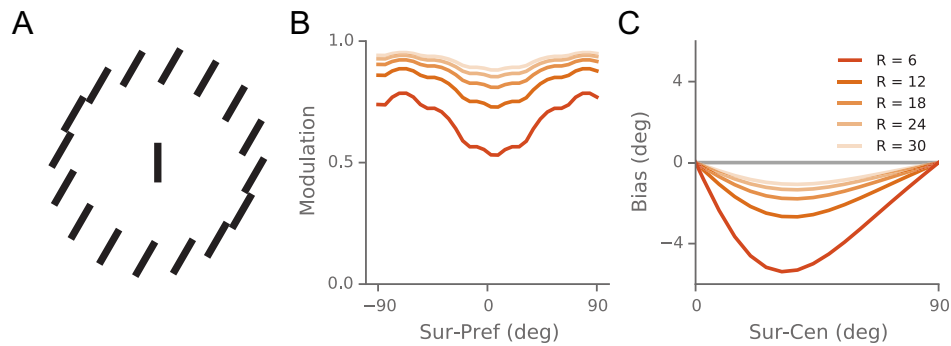


Fig. 6. Tilt illusion from elastica-based surround modulation by a grating type stimulus. (A) The stimulus consists of the center bar surrounded by a grating consisting of equally orientated bars. (B) The net modulation from the surrounding flankers across the population, relative to the surround orientation. As the orientation of the surrounding bars is varied relative the center orientation, the net response varies in a way reminiscent of surround modulation in V1, with strong suppression when the preferred orientation of a neuron and surround are aligned. (C) The bias is repulsive across all surround orientations.

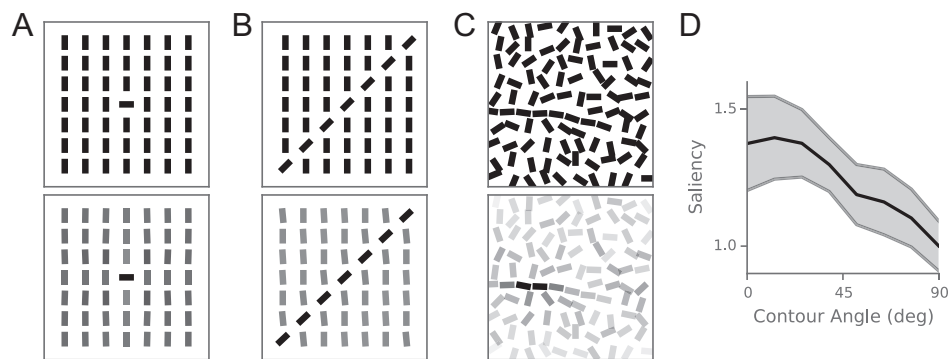


Fig. 7. Saliency effects resulting from elastica and contextual modulations. Top: input stimulus, bottom, perceived stimulus where the opacity represents the bar's saliency, s_b . To prevent edge effects, all scenes were rendered on a torus. (A) A single target stimulus pops out from the background, and does so more strongly based on how different the orientations are (now shown). The bars are placed on a grid such that the distance between two neighboring bars is 5. (B) Pop out of a simple contour in a uniform background. (C) Example of a noisy scene, when the angles between each bar in the contour are 11.25° . The spacing between the centers of each box in which a bar is randomly placed is 3. (D) How well the extraction works for different angles in the contour: as the angles are increased, contour saliency decreases, corresponding to the detection effects in Field et al. (1993). Gray error region indicates standard deviation of the mean.

enhance each others responses, resulting in the high saliency of the contour. Further note that the decoded bar orientations differ from the stimulus orientations.

We finally examine saliency of more general contours in random backgrounds. We use the method described in Field et al. (1993) to generate random images containing a random contour of length 8. Briefly, the contour is generated with a starting orientation and location, after which a set orientation change is made in a random direction (left or right), and a new bar is placed following the new orientation. An example stimulus containing a contour with orientation changes of 11.25° , is shown in Fig. 7C, with the decoded image in B. In Fig. 7D we quantify the contour saliency by calculating the average contour saliency over 50 different contour and background configurations, for different contour angle changes. As this angle increases, the saliency drops quickly, mirroring psychophysical contour detection probability (Field et al., 1993).

3.5. Dependency on model parameters

The elastica based contextual modulation has two parameters, E_0 and a , here we show that the essential features of the model do not depend on them. First, we fix $a = 0.1$ and vary E_0 to be 0, 4, or 8. Due to the fact that both contour extractions and the tilt illusion rely on only relative changes, they are fully invariant to changes in E_0 . However, the association field varies strongly with

the E_0 parameter, as would be expected from a parameter which mainly varies excitation versus inhibition, Fig. 8, top.

Next, we fix $E_0 = 1$, and set $a = 0.02, 0.1$ and 0.5 , Fig. 8, middle. As the gain parameter a changes the strength of the modulation, we see no change in the shape of the association field, but quantitative changes of the saliency of features and the decoding biases, displaying a trade-off between saliency strength and coding biases.

The increased bias with large a can be partially counteracted by narrowing the neural tuning width K_c (Eq. (4)), as contextual modulation results in a smaller shift of the population response when the tuning curves are sharp. Interestingly, the tuning width K_c also has an effect on the attraction effect in the tilt illusion with 6 flankers, Fig. 8, bottom row. As the tuning curve becomes sharper (i.e. larger K_c), the total illusion becomes weaker as expected, but the attractive reduces more. For $K_c = 1.5$ the attractive illusion completely disappears. This is because the attractive pull of the flankers is felt most strongly by neurons with a preferred orientation close to 90° . When the tuning curves are too narrow, these neurons do not respond, and thus the population response is not shifted.

4. Discussion

We have developed a computational model of V1 that implements the Gestalt Law of good continuation on a neural level through contextual modulations that were determined by the elastica energy. More specifically, the modulation by each bar outside a

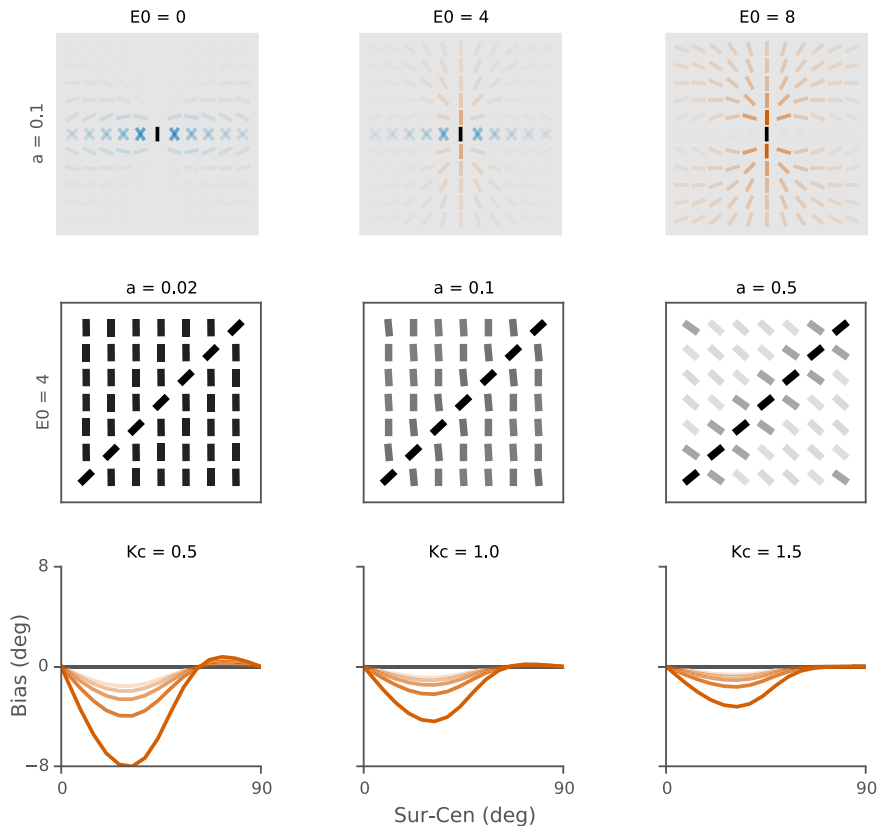


Fig. 8. Exploring the effect of the a , E_0 and K_c parameters in model. Top row: modulation strength a is kept constant, while E_0 is varied. The E_0 parameter determines the relative presence of excitation and inhibition in the association field. The illusions and saliency results depend on the relative difference of modulation and are not affected (not shown). Middle row: modulation offset E_0 is kept constant, while a is varied. While the relative presence of excitation and inhibition is not affected by this parameter, the strength of modulation is changed directly, resulting in larger differences in response. This results in stronger bias and saliency effects. Bottom row: neural tuning width K_c is varied while both E_0 and a are kept constant. While this generally affects the magnitude of the tilt illusions, the attractive illusion in the Westheimer experiment (Fig. 5F) disappears with large K_c .

neuron's receptive field is governed by the curvature energy of the smoothest curves connecting to it. This quite naturally lead to contour extraction, but more surprisingly also explains saliency detection, association fields, and various forms of the tilt illusion.

Our work builds on a large body of literature linking these various aspects of visual processing. Association fields have been derived from an image statistics perspective (Geisler et al., 2001; Sigman et al., 2001), and contour detection has been linked to association fields (Hansen & Neumann, 2008; Li, 1998, 1999; Spratling, 2012). Elastica and association fields have been linked before and used for contour detection and completion (Ernst et al., 2012; Sharon et al., 1997; Williams & Jacobs, 1997; Williams & Thorner, 2001). Finally, the tilt illusion for two flankers was explained from elastica in a Bayesian framework, but using a different expression for the modulation from us (Schwartz et al., 2006). Our model proposes a new explanation for several forms of the tilt illusions as following from individual elastica contour completions, including some counter-intuitive attractive effects. However, more importantly our model combines the phenomena described in earlier studies, both contour and illusion related, and links them together under the same basic elastica principle.

Not all aspects of the tilt illusion are captured by the model. Most prominently, the exact shape of the tilt illusion for collinear flankers as observed in Kapadia et al. (2000) was not reproduced, Fig. 5E. However, we did find attraction for smaller angles, and repulsion otherwise, which has not been explained before. The attraction disappears as the flankers are placed further away, also in accordance with what was found in Kapadia's work. An important factor in producing these effects is calculating the

elastica curves from the centers of the bars. Although the flanker ends were level with the center bar, this allowed for both the attractive and repulsive effects dependent on orientation and distance. This might suggest that neurons with a receptive field of smaller scale than the flanker bar drive the attractive effect in humans. However, such an explanation would require assumptions about how tilt estimates at different spatial scales are combined.

Interestingly, the illusion caused by a hexagon of 6 flankers was captured by the model, with both a repulsive and attractive effect. This effect was dependent on the tuning to the center bar. Although the attraction completely disappeared for very narrow tuning $K_c = 1.5$, most neurons have broad tuning corresponding to $K_c \approx 0.5$ (Cavanaugh, Bair, & Movshon, 2002). Next, when in order to mimic gratings we increased the number of flankers and rotated the entire surround, the attractive effect disappeared for more than 8 flankers (for our parameters). However the repulsive illusion was always present. Existing explanations of the tilt illusion have taken either a mechanistic view, where it is purely a result of the surround suppression (Clifford et al., 2000), or functional views such as arising from image statistics (Schwartz, Sejnowski, & Dayan, 2009), Bayesian processing (Schwartz et al., 2006), or as some form of image normalization (Clifford, 2014). We propose a new hypothesis: the smoothest continuations of the surround elements tilts the percept away from the surround and, in special cases, attracts it.

Despite the model's simplicity, we consider the model biologically feasible. The contextual modulations are effected as independent contributions from each flanker, as one would expect for modulation from individual surround neurons. The resulting

modulation matches electrophysiology; both for individual flanker contributions in the form of the neural association field, with excitatory effects for collinearity and inhibitory effects for parallel bars (Kapadia et al., 2000; Kinoshita et al., 2009), and the net effect for many surrounding bars which leads to suppressive surround modulation (Cavanaugh et al., 2002; Gilbert & Wiesel, 1990; Seriès, Lorenceau, & Frégnac, 2003). Because the modulation relies on pairwise interactions only, it is plausible that some form of Hebbian learning shapes its tuning (Bednar, 2012). A caveat is that the statistics of natural images which include both textures and contours are dominated by parallel structures; an association field arises only when the statistics are restricted to contours (Geisler et al., 2001).

Currently we have assumed a single contrast level, which is clearly unrealistic for most natural images. In particular, association fields are known to change with contrast. Low contrast leads predominantly to excitation, while high contrast leads predominantly to inhibition (Cavanaugh et al., 2002; Kapadia et al., 2000). It is possible to extend the model to describe responses to stimuli with heterogeneous contrasts. The center contrast can be represented by the A_c parameter (e.g. Sclar, Maunsell, & Lennie, 1990). The contrast of the flanker can be coded in the a and E_0 parameters of each modulation term. In particular, changing E_0 as in Fig. 8 top row, qualitatively matches the observed contrast dependence of the association fields in Kapadia et al.

It is perhaps not surprising that elastica models the contextual interactions of V1 well, if the interactions do indeed exist for the purpose of detecting contours. Besides the elastica curves being especially pleasing to the eye, contours of natural objects are often well described by elastica curves. As an example in this paper, the shape in Fig. 2B would be a good candidate for a leaf. This is the very reason it is used in computer vision for contour completion of partially hidden objects (e.g. Kimia, Frankel, & Popescu, 2003; Mumford, 1994; Zhou, Zheng, & Yang, 2012).

Apart from the illusions of Fig. 5A, C and D, our model makes several predictions: first, the inhibitory connections seem to be broader tuned than excitatory connections. However, which inhibitory connections are the strongest is strongly dependent on relative position and orientation, Fig. 4. This could be tested experimentally. Secondly, our contextual modulations affect a neuron individually, and lead to contextual interactions both for small sets of bars and full surrounds. Experimentally, neither neurophysiologically nor psychophysically, it is known if these are linked. I.e., it is unknown whether the modulation by a surround built up with individual elements, can be explained from its individual contributions.

The neural character of the model allows for a number of straightforward extensions: (1) it will be interesting to include more realistic Gabor-type receptive fields at a variety of scales. (2) Currently the bars are assumed to have zero length. It is straightforward to find the elastica curves connecting the ends of the bars, with the only minor complication that β_f and β_c now become dependent on θ_f and θ_c . However, without a more realistic receptive field such an extension is rather ad hoc. (3) In the current implementation flankers modulate the center, but there is no recurrent feedback in which the modulated response change the activity of the flankers. In this sense the model performs a one-step approximation, which is valid as long as the shifts in the tuning curve are moderate. In a recurrent model, the dynamics of the illusions presented here would be of interest.

Acknowledgements

We thank Matthew Graham and James A. Bednar for discussions. S.W. Keemink was supported by the EuroSpin Erasmus

Mundus program and the EPSRC Doctoral Training Centre in Neuroinformatics (EP/F500385/1 and BB/F529254/1).

References

- Bauer, R., & Heinze, S. (2002). Contour integration in striate cortex. *Experimental Brain Research*, 147(2), 145–152.
- Bednar, J. (2012). Building a mechanistic model of the development and function of the primary visual cortex. *Journal of Physiology*, 106, 194–211.
- Bosking, W. H., Zhang, Y., Schofield, B., & Fitzpatrick, D. (1997). Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *The Journal of Neuroscience*, 17(6), 2112–2127.
- Bruckstein, A. M., & Netravali, A. N. (1990). On minimal energy trajectories. *Computer Vision, Graphics, and Image Processing*, 49(3), 283–296.
- Cavanaugh, J. R., Bair, W., & Movshon, J. (2002). Selectivity and spatial distribution of signals from the receptive field surround in macaque V1 neurons. *Journal of Neurophysiology*, 88, 2547–2556.
- Cavanaugh, J. R., Bair, W., & Movshon, J. (2002). Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons. *Journal of Neurophysiology*, 88, 2530–2546.
- Clifford, C. W. G., Wenderoth, P., & Spehar, B. (2000). A functional angle on some after-effects in cortical vision. *Proceedings of the Royal Society B: Biological Sciences*, 267(1454), 1705–1710.
- Clifford, C. (2014). The tilt illusion: Phenomenology and functional implications. *Vision Research*, 104, 3–11.
- Ernst, U., Mandon, S., Schinkel-Bielefeld, N., Neitzel, S., Kreiter, A., & Pawelzik, K. (2012). Optimality of human contour integration. *PLoS Computational Biology*, 8(5), e1002520.
- Field, D., Hayes, A., & Hess, R. (1993). Contour integration by the human visual system evidence for a local association field. *Vision Research*, 33(2), 1117–1131.
- Geisler, W., Perry, J., Super, B., & Gallogly, D. (2001). Edge co-occurrence in natural images predicts contour grouping and performance. *Vision Research*, 41, 711–724.
- Georgopoulos, A., Schwartz, A., & Kettner, R. (1986). Neuronal population coding of movement direction. *Science*, 233(4771), 1416–1419.
- Gilbert, C., & Wiesel, T. (1990). The influence of contextual stimuli on the orientation selectivity of cells in primary visual cortex of the cat. *Vision Research*, 30, 1689–1701.
- Hansen, T., & Neumann, N. (2008). A recurrent model of contour integration in primary visual cortex. *Journal of Vision*, 8, 1–25.
- Kapadia, M., Ito, M., Gilbert, C., & 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.
- Kapadia, M., Westheimer, G., & Gilbert, C. (2000). Spatial distribution of contextual interactions in primary visual cortex and in visual perception. *Journal of Neurophysiology*, 84(4), 2048–2062.
- Kimia, B., Frankel, I., & Popescu, A. (2003). Euler spiral for shape completion. *International Journal of Computer Vision*, 54(1/2/3), 159–182.
- Kinoshita, M., Gilbert, C., & Das, A. (2009). Optical imaging of contextual interactions in V1 of the behaving monkey. *Journal of Neurophysiology*, 102, 1930–1944.
- Koffka, K. (1935). *Principles of gestalt psychology*. London: Lund Humphries.
- Leung, T., & Malik, J. (2001). Contour continuity in region based image segmentation. *International Journal of Computer Vision*, 43(1), 7–27.
- Levien, R. (2008). The elastica: A mathematical history. Technical report UCB/Eecs-2008-103. Berkeley: Eecs Department.
- Li, W., Piëch, V., & Gilbert, C. (2006). Contour saliency in primary visual cortex. *Neuron*, 50, 951–962.
- Li, Z. (1998). A neural model of contour integration in the primary visual cortex. *Neural Computation*, 10, 903–940.
- Li, Z. (1999). Contextual influences in V1 and as a basis for pop out and asymmetry and in visual search. *Proceedings of the National Academy of Sciences of the United States of America*, 96, 10530–10535.
- Mumford, D. (1994). Elastica and computer vision. In C. Bajaj (Ed.), *Algebraic geometry and its applications* (pp. 491). Springer.
- Nothdurft, H. (1993). The conspicuousness of orientation and motion contrast. *Spatial Vision*, 7, 341–363.
- Schwartz, O., Sejnowski, T. J., & Dayan, P. (2006). A Bayesian framework for tilt perception and confidence. *Advances in Neural Information Processing Systems*, 18, 2680–2718.
- Schwartz, O., Sejnowski, T. J., & Dayan, P. (2009). Perceptual organization in the tilt illusion. *Journal of Vision*, 9(4), 1–20.
- Schwartz, O., Hsu, A., & Dayan, P. (2007). Space and time in visual context. *Nature Reviews Neuroscience*, 8(7), 522–535.
- Sclar, G., Maunsell, J. H. R., & Lennie, P. (1990). Coding of image contrast in the central visual pathways of the macaque monkey. *Vision Research*, 30, 1–10.
- Seriès, P., Lorenceau, J., & Frégnac, Y. (2003). The silent surround of V1 receptive fields: Theory and experiments. *Journal of Physiology-Paris*, 97(4–6), 453–474.
- Sharon, E., Brandt, A., & Basri, R. (1997). Completion energies and scale. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 10(10), 1117–1131.
- Shushruth, S., Nurminen, L., Bijanzadeh, M., Ichida, J. M., Vanni, S., & Angelucci, A. (2013). Different orientation tuning of near- and far-surround suppression in macaque primary visual cortex mirrors their tuning in human perception. *Journal of Neuroscience*, 33(1), 106–119.

- Sigman, M., Cecchi, G., Gilbert, C., & Magnasco, M. (2001). On a common circle: Natural scenes and gestalt rules. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(4), 1935–1940.
- Spratling, M. (2012). Predictive coding as a model of the V1 saliency map hypothesis. *Neural Networks*, *26*, 7–28.
- Tang, Q., Sang, N., & Zhang, T. (2007). Extraction of salient contours from cluttered scenes. *Pattern Recognition*, *40*, 3100–3109.
- Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt: II. *Psychologische Forschung*, *4*, 301–350.
- Westheimer, G. (1990). Simultaneous orientation contrast for lines in the human fovea. *Vision Research*, *30*(11), 1913–1921.
- Williams, L., & Jacobs, D. (1997). Stochastic completion fields: A neural model of illusory contour shape and salience. *Neural Computation*, *9*(4), 837–858.
- Williams, L., & Thornber, K. (2001). Orientation, scale, and discontinuity as emergent properties of illusory contour shape. *Neural Computation*, *13*, 1683–1711.
- Zhou, H., Zheng, J., & Yang, X. (2012). Euler arc splines for curve completion. *Computers & Graphics*, *36*, 642–650.