ORIGINAL PAPER

A CORF computational model of a simple cell that relies on LGN input outperforms the Gabor function model

George Azzopardi · Nicolai Petkov

Abstract Simple cells in primary visual cortex are believed to extract local contour information from a visual scene. The 2D Gabor function (GF) model has gained particular popularity as a computational model of a simple cell. However, it short-cuts the LGN, it cannot reproduce a number of properties of real simple cells, and its effectiveness in contour detection tasks has never been compared with the effectiveness of alternative models. We propose a computational model that uses as afferent inputs the responses of model LGN cells with center–surround receptive fields (RFs) and we refer to it as a Combination of Receptive Fields (CORF) model. We use shifted gratings as test stimuli and simulated reverse correlation to explore the nature of the proposed model. We study its behavior regarding the effect of contrast on its response and orientation bandwidth as well as the effect of an orthogonal mask on the response to an optimally oriented stimulus. We also evaluate and compare the performances of the CORF and GF models regarding contour detection, using two public data sets of images of natural scenes with associated contour ground truths. The RF map of the proposed CORF model, determined with simulated reverse correlation, can be divided in elongated excitatory and inhibitory regions typical of simple cells. The modulated response to shifted gratings that this model shows is also characteristic of a simple cell. Furthermore, the CORF model exhibits cross orientation suppression, contrast invariant orientation tuning and response saturation. These properties are observed in real simple cells, but are not possessed by the GF model. The proposed CORF model outperforms the GF model in contour detection with high statistical confidence (RuG data set: $p < 10^{-4}$, and Berkeley data set: $p < 10^{-4}$). The proposed CORF model is more realistic than the GF model and is more effective in contour detection, which is assumed to be the primary biological role of simple cells.

1 Introduction

The majority of neurons in primary visual cortex (area V1) exhibit orientation selectivity (Hubel and Wiesel 1962). Typically, such a neuron would respond to an edge or a line of a given orientation in a given area of the visual field, called its receptive field (RF). Hubel and Wiesel identified two main classes of neuron that they called simple and complex cells. The RFs of simple cells can be divided in excitatory and inhibitory regions while no such division is possible in complex cells. These findings gave rise to an active area of research (Hubel and Wiesel 1974; Macleod and Rosenfeld 1974; Tyler 1978; De Valois et al. 1978, 1979; Albrecht et al. 1980; Von Der Heydt 1987) and the pioneers Hubel and Wiesel were later awarded a Nobel prize (Hubel 1982).

The above electrophysiological findings had instigated the development of various computational models. Simple cells are typically modeled by linear spatial summation followed by half-wave rectification (Movshon et al. 1978a; Andrews and Pollen 1979; Maffei et al. 1979; Glezer et al. 1980; Kulikowski and Bishop 1981; Daugman 1985; Jones and Palmer 1987). Marcelja (1980) has suggested that the elementary one-dimensional signals studied by Gabor (1946) can be used to model the structure of the RFs of simple cells. Later Daugman (1985) extended this idea to a family of two-dimensional (2D) Gabor functions (GFs) that have

G. Azzopardi (✉) · N. Petkov
Johann Bernoulli Institute for Mathematics and Computer Science,
University of Groningen, Groningen, The Netherlands
e-mail: g.azzopardi@rug.nl

been claimed to fit well the 2D RF profiles of cat simple cells (Jones and Palmer 1987). The validity of GFs for modeling of simple cells has also been questioned (Stork and Wilson 1990) and alternative functions have been proposed, such as differences of offset Gaussians.

The GF model does not take into account the anatomical structure of the visual system as it uses as inputs the intensity values of an image as these are projected on the retina and bypasses the lateral geniculate nucleus (LGN) within the thalamus (Chung and Ferster 1998; Ferster et al. 1996; Reid and Alonso 1995). Furthermore, it fails to reproduce some properties of simple cells, such as cross orientation suppression, independence of orientation tuning on contrast, and response saturation.

Hubel and Wiesel (1962) speculated that the RF profile of a simple cell and its orientation selectivity are the result of the specific alignment of the RFs of LGN cells that provide input. In this study, we consider such a computational model of a simple cell and call it the Combination of Receptive Fields (CORF) model.

The biological role of orientation-selective cells is believed to be the extraction of local contour information, which is a fundamental step for further, more complex visual tasks, such as object recognition (Morrone and Owens 1987; Morrone and Burr 1988; Rosenthaler et al. 1992; Mehrotra et al. 1992; Heitger 1995; Kovesi 1999). The performance of various computational models of a simple cell in contour detection tasks has, however, not been quantified and they have not been compared in that respect. In the following, we evaluate and compare the performances of the proposed CORF model and the GF model, using two public data sets of images of natural scenes with associated contour ground truths. We also compare their ability to reproduce other properties of simple cells, such as cross orientation suppression, independence of orientation tuning on contrast, and response saturation.

The paper is organized as follows. In Sect. 2, we present the CORF model. In Sect. 3, we demonstrate that it is a model of a simple cell and compare it with the GF model in different respects. Section 4 contains a discussion of some aspects of the model. We give a brief summary in Sect. 5 and draw conclusions in Sect. 6. We explain simulated reverse correlation in Appendix A. In Appendix B, we elaborate on the choice of an output function for the proposed model and in Appendix C, we give details about the evaluation method of contour localization that we use in our experiments.

2 Computational model

2.1 Overview

Figure 1 illustrates the general set-up of the proposed CORF computational model. Each of the light and dark disks in
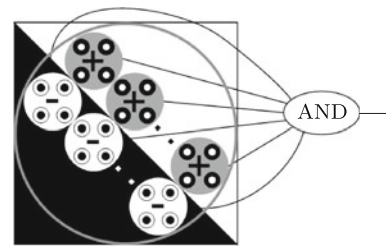
Fig. 1 Sketch of the proposed CORF computational model of a simple cell. Orientation selectivity is achieved by combining the responses of two parallel sets of co-linear sub-units: one of center-on ("+") and the other of center-off ("−") type. Each sub-unit computes a sum of the weighted responses of a local group of model LGN cells

Fig. 1 represents the RF of a sub-unit that receives input from a pool of center-on ("+") or center-off ("−") model LGN cells. Such model LGN cells detect contrast changes. A sub-unit computes the sum of the weighted responses of the model LGN cells it receives input from. These model LGN cells have the same polarity (on or off) and RF size and neighboring RFs. The RF of the sub-unit is the union of the RFs of the involved model LGN cells and it has the same polarity (on or off) as these cells. In this way, a sub-unit detects contrast changes, similar to a model LGN cell, but it does so in a wider area. A sub-unit can be thought of as a dendrite branch of a simple cell which receives input from a pool of adjacent LGN cells.

The orientation selectivity of a CORF model cell is achieved by combining the responses of given sub-units with appropriate polarities and alignment of their RFs, as illustrated in Fig. 1. The model parameters are determined in an automatic configuration process in which an example edge of a given orientation and polarity is presented. This input stimulus gives rise to a certain local configuration of model LGN cell activities in the RF of the concerned CORF model cell. This local configuration is used to determine the polarity of the involved sub-units and their mutual spatial arrangement. The response of the considered CORF model cell is computed as the weighted geometric mean of the sub-unit responses. In this way, a CORF model cell generates a response only when all its afferent inputs are stimulated. In the following sections, we explain the model configuration process in more detail.

2.2 Model LGN cells with center–surround RFs

We use a difference of 2D Gaussian functions to model an LGN cell (Rodieck 1965; Croner and Kaplan 1995):

$$\mathrm{DoG}_\sigma^+(x, y) \stackrel{\mathrm{def}}{=} \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \tag{1}$$
$$- \frac{1}{2\pi(0.5\sigma)^2} \exp\left(-\frac{x^2 + y^2}{2(0.5\sigma)^2}\right)$$
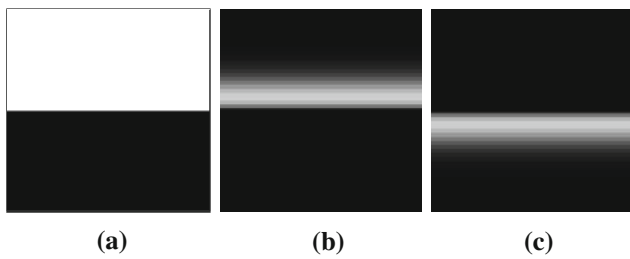
Fig. 2 **a** Synthetic input image of a sharp edge. Responses of **b** center-on and **c** center-off model LGN cells. The response of a model LGN cell is rendered as the brightness of a pixel on which the RF of that cell is centered

Fig. 3 Configuration of a CORF model cell. **a** The *largest circle* depicts the RF boundary of the model, the center of which is illustrated by the "+" marker and lies on a vertical edge. **b** The *top* and *bottom* plots show the responses of model LGN cells along the *outer* and *inner* interrupted *circles* in (**a**). The *labeled small spots* represent the RF centers of eight sub-units, four center-on (*white spots* with *black* boundary) and four center-off (*black spots* with *white* boundary) with coordinates that are determined by the corresponding labeled local maxima in the considered responses of model LGN cells.

where σ is the standard deviation of the outer Gaussian function. We fix the standard deviation of the inner Gaussian function to 0.5σ, which is in accordance with electrophysiological findings of LGN cells in mammals (Irvin et al. 1993; Xu et al. 2002). A $\mathrm{DoG}_\sigma^+(x, y)$ function corresponds to a center-on RF, such that the central region is excitatory and the surrounding is inhibitory. A center-off RF is denoted by $\mathrm{DoG}_\sigma^-(x, y)$ and is defined as the negative of $\mathrm{DoG}_\sigma^+(x, y)$:

$$\mathrm{DoG}_\sigma^-(x, y) \stackrel{\mathrm{def}}{=} -\mathrm{DoG}_\sigma^+(x, y) \tag{2}$$

The response of a model LGN cell with a RF centered at image coordinates (x, y) is computed by linear spatial summation of the intensity distribution $I(x', y')$ in the input image, weighted with the function $\mathrm{DoG}(x - x', y - y')$, followed by half-wave rectification[1]:

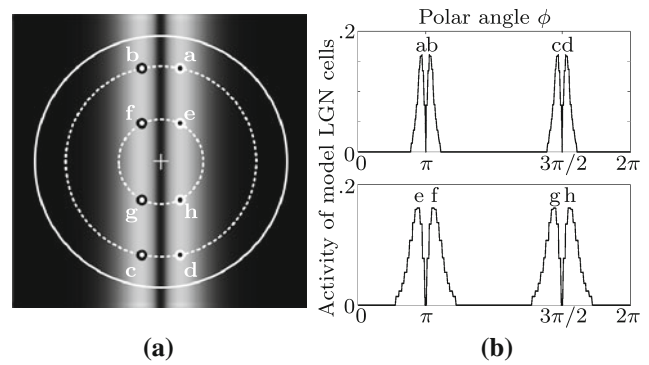$$c_\sigma^\delta(x, y) \stackrel{\mathrm{def}}{=} |I \star \mathrm{DoG}_\sigma^\delta|^+ \tag{3}$$

where the symbol δ represents the polarity (+ for center-on and − for center-off) of the DoG function used.

Figure 2a shows a synthetic input image of a sharp step edge and in Fig. 2b and c, we show the corresponding responses of center-on and center-off model LGN cells.

2.3 Sub-unit and its parameters

Figure 3 illustrates the process of configuring a CORF model cell. The outer circle in Fig. 3a demarks the RF of that cell. Its center is positioned on an edge (in the input image) which gives rise to model LGN cell responses rendered in Fig. 3a. The eight small spots represent the RF centers of eight sub-units, four of center-on and four of center-off type. The inclusion of such sub-units in the computational model is decided on as follows. We take two (in general k) concentric circles centered on the RF center and consider the responses of model LGN cells along these circles, Fig. 3b. The positions along these circles at which these responses reach significant

local maxima are the positions at which we include sub-units. For the considered example, there are four such positions for each of the two circles, which results in the inclusion of eight sub-units in the model at hand. The number of sub-units depends on the number of circles we consider and the specific input pattern presented at the time of configuration. In our model, each sub-unit included is represented in parametric form by a tuple $(\delta, \sigma, \rho, \phi)$ where the parameters represent the polarity δ of the sub-unit, the scale parameter σ of the involved model LGN cells in its pool, the radius ρ and the polar angle ϕ of the RF center of the sub-unit relative to the RF center of the CORF model cell, respectively.

We denote by $S = \{(\delta_i, \sigma_i, \rho_i, \phi_i) | i = 1 \dots n\}$ the set of 4-tuples that represent the configured sub-units above. For the concerned model in Fig. 3, for $\sigma = 5$, and two configuration circles ($\rho \in \{18, 34\}$) for an image of size 100×100 pixels, the method described above results in eight sub-units, which are specified by the tuples in the following set:

$S = \{$

$(\delta_1 = -, \sigma_1 = 5, \rho_1 = 34, \phi_1 = 1.48),$

$(\delta_2 = +, \sigma_2 = 5, \rho_2 = 34, \phi_2 = 1.66),$

$(\delta_3 = +, \sigma_3 = 5, \rho_3 = 34, \phi_3 = 4.62),$

$(\delta_4 = -, \sigma_4 = 5, \rho_4 = 34, \phi_4 = 4.80)$

$(\delta_5 = -, \sigma_5 = 5, \rho_5 = 18, \phi_5 = 1.41),$

$(\delta_6 = +, \sigma_6 = 5, \rho_6 = 18, \phi_6 = 1.74),$

$(\delta_7 = +, \sigma_7 = 5, \rho_7 = 18, \phi_7 = 4.55),$

$(\delta_8 = -, \sigma_8 = 5, \rho_8 = 18, \phi_8 = 4.88)$

$\}$

---

[1] We use $\star$ and $|.|^+$ to denote convolution and half-wave rectification, respectively.

The first tuple in the set S, for instance, describes a sub-unit that collects its afferent inputs from the responses of center-off ($\delta_1 = -$) model LGN cells with a RF size characterized by a standard deviation of ($\sigma_1 =$) 5 pixels, around a position at a radius of ($\rho_1 =$) 34 pixels and an angle of ($\phi_1 =$) 1.48 radians with respect to the RF center of the CORF model cell; the RF center of this sub-unit is marked by "a" in Fig. 3a.

2.4 Sub-unit responses

We denote by $s_{\delta_i,\sigma_i,\rho_i,\phi_i}(x, y)$ the response of a sub-unit, which we compute by linear spatial summation of the half-wave rectified responses $c_{\sigma_i}^{\delta_i}(x, y)$ of model LGN cells with preferred polarity δ_i and scale σ_i around position (ρ_i, ϕ_i) with respect to the RF center of the CORF model cell, weighted with a 2D Gaussian function $G_{\sigma'}$:

$$s_{\delta_i,\sigma_i,\rho_i,\phi_i}(x, y) \overset{\text{def}}{=} \tag{4}$$
$$\sum_{x'} \sum_{y'} \{c_{\sigma_i}^{\delta_i}(x - \Delta x_i - x', y - \Delta y_i - y')G_{\sigma'}(x', y')\}$$
$$\Delta x_i = -\rho_i \cos \phi_i, \quad \Delta y_i = -\rho_i \sin \phi_i,$$
$$-3\sigma' \le x', \quad y' \le 3\sigma'$$

The standard deviation σ' of the Gaussian function $G_{\sigma'}$ is a linear function[2] of the parameter ρ which is consistent with neurophysiological evidence for the relationship between the eccentricity and the average RF diameter of LGN cells (Xu et al. 2001). Equation 4 presents a convolution of the weighting function $G_{\sigma'}$ with the function $c_{\sigma_i}^{\delta_i}(x, y)$ that is shifted by ($\Delta x_i, \Delta y_i$) where this shift vector is determined by the sub-unit parameters (ρ_i, ϕ_i).

### 2.5 Combining sub-unit responses

We define the response r_S of a CORF model cell as the weighted geometric mean of all sub-unit responses that belong to the specific selection determined by the set S:

$$r_S(x, y) \overset{\text{def}}{=} \left(\prod_{i=1}^{|S|} \left(s_{\delta_i,\sigma_i,\rho_i,\phi_i}(x, y) \right)^{\omega_i} \right)^{1/\sum_{i=1}^{|S|} \omega_i} \tag{5}$$
$$\omega_i = \exp^{-\frac{\rho_i^2}{2\sigma'^2}}, \quad \sigma' = \frac{1}{3} \max_{i \in \{1...|S|\}} \{\rho_i\}$$

Computing the product of sub-unit responses means that the concerned CORF model cell is activated only when all its afferent sub-units are active. The input contribution of sub-units decreases with an increasing distance of their RF centers from the RF center of the CORF model cell.

[2] The standard deviation σ' of the 2D Gaussian function $G_{\sigma'}$ is computed as $\sigma' = (d_0 + \alpha\rho)/6$. We set $d_0 = 2$ and $\alpha = 0.9$.
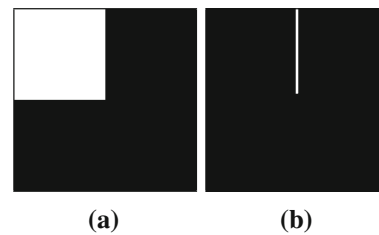
Fig. 4 a A synthetic image (512×512 pixels) and **b** the corresponding responses of a CORF model that is selective for vertical edges ($r_S(x, y)$)
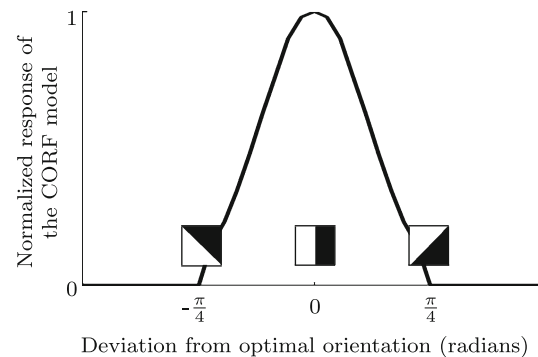
Fig. 5 Orientation selectivity of the CORF model

As illustrated in Fig. 4 the configured CORF model cell responds to a vertical edge in the given synthetic image.

Figure 5 illustrates the orientation selectivity of the concerned model. A maximum response is obtained for the orientation for which the model was configured. The response declines with the deviation of the orientation of the input stimulus from the optimal one and practically disappears when this deviation is greater than $\pi/4$ radians. Qualitatively, this is in line with the orientation tuning recorded by De Valois et al. (1982) where the majority of simple cells in macaque visual cortex have an orientation bandwidth of 0.7 radians at half amplitude.

2.6 Treating different orientations

The orientation preference of a CORF model cell as defined above depends on the orientation of the edge used for the model configuration. One can create models with different orientation preferences by presenting different edges. Alternatively, one can manipulate the parameters in the set S, which corresponds to orientation preference for 0° to obtain a new set $\Re_\psi(S)$ with orientation preference ψ. We define $\Re_\psi(S)$ as follows:

$$\Re_\psi(S) \overset{\text{def}}{=} \{(\delta_i, \sigma_i, \rho_i, \phi_i + \psi) | \forall (\delta_i, \sigma_i, \rho_i, \phi_i) \in S\} \tag{6}$$

For the detection of contours of any orientation, we merge the responses of CORF models with different orientation preferences by taking the maximum value at a given location (x, y):

(a)　　　　　　　**(b)**

(c)　　**(d)**　　**(e)**　　**(f)**

(g)　　**(h)**　　**(i)**　　**(j)**

(k)　　**(l)**　　**(m)**　　**(n)**
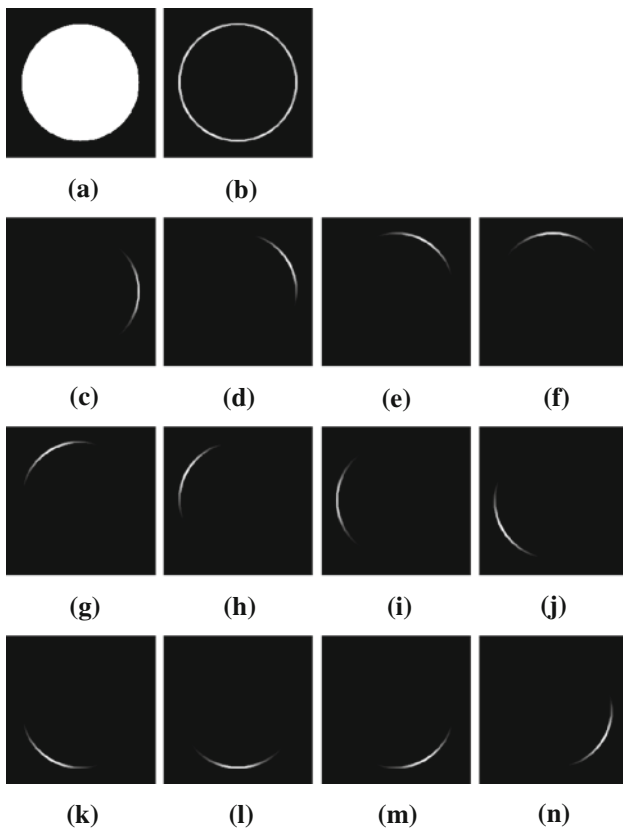
Fig. 6 **a** Synthetic input image (512×512 pixels). **b** The maximum superposition ($\hat{r}_S(x, y)$) of the **c**–**n** responses of CORF models that are selective for 12 equidistant orientations. **c** $\psi = 0$, **d** $\psi = \frac{\pi}{6}$, **e** $\psi = \frac{\pi}{3}$, **f** $\psi = \frac{\pi}{2}$, **g** $\psi = \frac{2\pi}{6}$, **h** $\psi = \frac{5\pi}{6}$, **i** $\psi = \pi$, **j** $\psi = \frac{7\pi}{6}$, **k** $\psi = \frac{4\pi}{3}$, **l** $\psi = \frac{3\pi}{2}$, **m** $\psi = \frac{5\pi}{3}$, **n** $\psi = \frac{11\pi}{6}$

$$\hat{r}_S(x, y) \stackrel{\text{def}}{=} \max_{\psi \in \Psi} \{r_{\Re_\psi(S)}(x, y)\} \tag{7}$$

where Ψ is a set of n_θ equidistant orientations given as $\Psi = \{\frac{2\pi}{n_\theta} i | 0 \le i < n_\theta\}$. Figure 6a illustrates a synthetic input image of a bright disk. Figure 6b shows the maximum superposition $\hat{r}_S(x, y)$ of the responses, shown in Fig. 6c–n, of a set of CORF models that are selective for ($n_\theta =$) 12 equidistant orientations. As demonstrated in Fig. 6b, the choice of $n_\theta = 12$ ensures sufficient response for all orientations.

3 Results: CORF versus GF model comparison

3.1 Is CORF a simple cell model?

A classical means of testing an orientation selective cell for its type, simple versus complex, is to record its response to a grating of its preferred orientation and spatial frequency and observe if the response changes when the grating is shifted. A complex cell shows unmodulated response while the response of a simple cell alternates between high and low

$\cos(\omega x + \psi)$

(a)　　　　　　　**(b)**
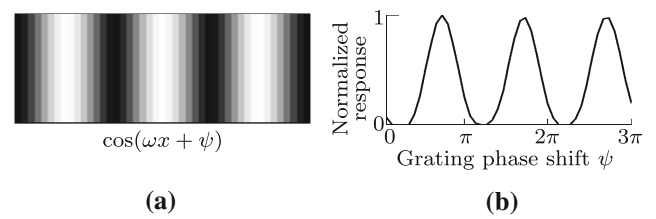
Fig. 7 **a** Vertical sinusoidal grating ($\cos(\omega x + \psi)$) and **b** normalized response (in a given point) of a CORF model cell to this grating as a function of the phase offset ψ. The response of the model alternates between high and low values with a changing phase offset ψ (that corresponds to shifting the grating) as the response of a real simple cell would do


(a)　　　　**(b)**　　　　**(c)**
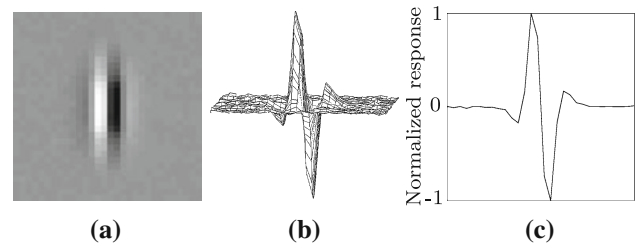
Fig. 8 RF map of a CORF model cell determined with simulated reverse correlation, rendered as **a** intensity map and **b** 2D function. **c** One-dimensional profile of the function along a RF cross section that is perpendicular to the preferred orientation

values with the grating shift. We computed the responses of the concerned CORF model cell to a grating with a changing spatial phase offset and observed that this response is qualitatively typical of the behavior of a simple cell, Fig. 7.

An important feature of a simple cell that distinguishes it from a complex cell is that its RF can be divided into excitatory and inhibitory regions. Since the CORF model cell that we propose above computes the response as a product of weighted LGN responses, it is not clear that its RF can be divided in such regions. To explore this aspect of the model, we apply to it *simulated* reverse correlation as explained in detail in Appendix A. Figure 8 shows a map determined in this way for a CORF model cell. The map exhibits two elongated regions: an excitatory one (rendered bright) in which the values are positive and an inhibitory one (shown dark) in which the values are negative, characteristic of the RF profiles of real simple cells.

We should point out that the map shown in Fig. 8 is not an impulse response that can be used for weighted summation of the input stimulus followed by half-wave rectification or thresholding, as this is the case with the GF model. The GF model and any other semi-linear model based on weighted spatial summation will produce a non-zero response that grows with stimulus contrast for any edge orientation but the one that is strictly orthogonal to the preferred orientation. This is a drawback of such models because it is in contrast to neurophysiological evidence that the orientation bandwidth
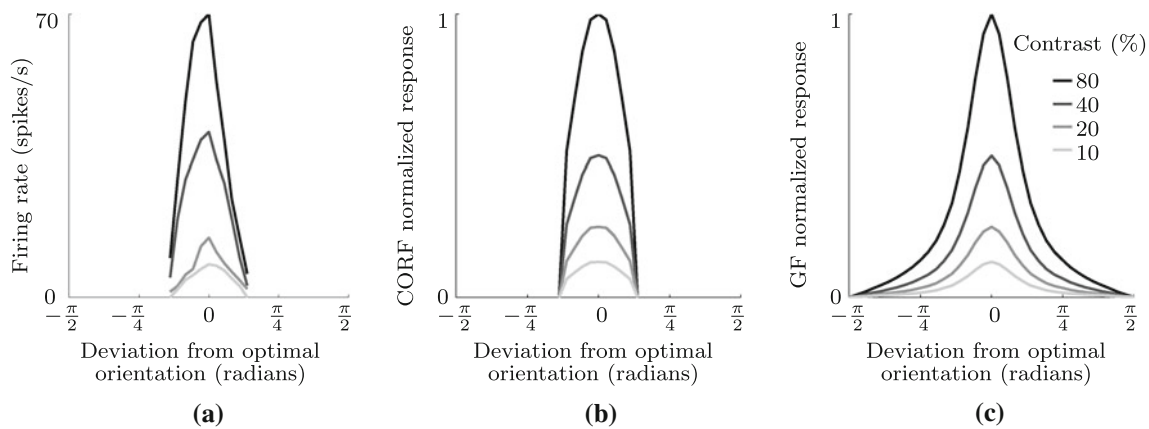
Fig. 9 Orientation tuning curves at four different contrasts obtained for **a** a real simple cell (redrawn from Sclar and Freeman 1982), **b** a CORF model cell with similar bandwidth, and **c** a GF model. The GF model gives response at any stimulus orientation but the one that is strictly orthogonal to the preferred orientation. If a threshold is used, the orientation bandwidth of the GF model becomes dependent on stimulus contrast, which is not the case with real simple cells


Fig. 10 Model LGN cells with RFs that fall in homogeneous areas of the stimulus provide zero input and therefore the response of the CORF model cell is zero. This property causes the CORF model cell to have a band-limited orientation tuning curve
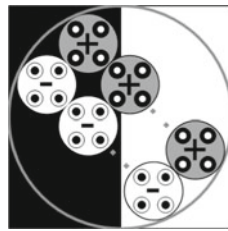
of simple cells is not affected by the contrast of the stimuli, Fig. 9a (see e.g., Finn et al. 2007 and references therein). The orientation bandwidth of the CORF model cell can be controlled by appropriate selection of the model parameters, such as the given ρ and σ values. Similar to real simple cells, the resulting orientation bandwidth of the CORF model cell is invariant to contrast, Fig. 9b. On the other hand, the orientation bandwidth of the GF model can only be constrained using a threshold, which will lead to a broadening orientation tuning with increasing contrasts (the so-called iceberg effect), Fig. 9c.

The key to understanding the orientation tuning of the CORF model cell is its multiplicative character: the response is computed as a *product* of the afferent inputs and if any of these is zero, the response will be zero, independent of how strong the other inputs are. For instance, the CORF model cell, the structure of which is depicted in Fig. 10, gives a zero response to an edge stimulus of high contrast with an orientation that differs substantially from the preferred orientation due to the zero input that it receives from some of the afferent groups (sub-units) of model LGN cells.

Another property of real simple cells, that cannot be reproduced by semi-linear spatial summation models such as the GF model, is a cross orientation suppression (Priebe and Ferster 2006). This means that an oriented mask stimulus that is orthogonal to a simultaneously presented test stimulus of the
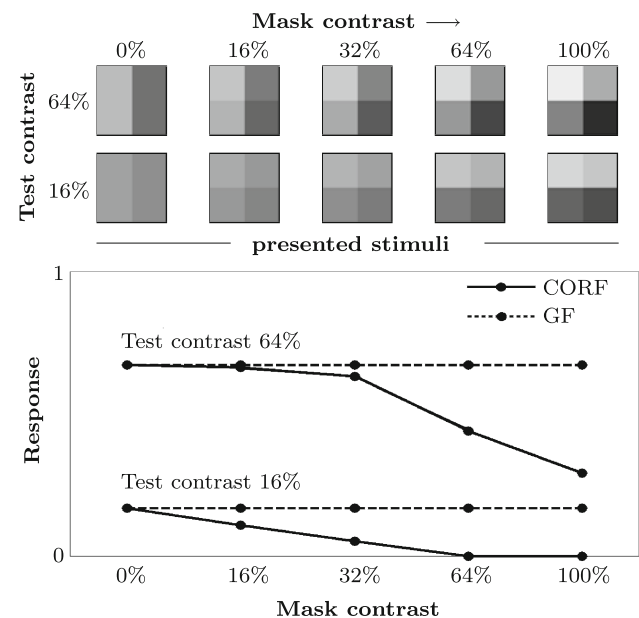
Fig. 11 The CORF model cell exhibits cross orientation suppression, while the GF model does not. The plots illustrate the responses of the CORF and the GF models to the images shown at the *top*. In these experiments, the RFs of the models are placed in the *center* of a given image. Every image is generated by superimposing a horizontal-edge stimulus (mask) on a vertical-edge stimulus (test) which is the preferred orientation stimulus of the concerned models. Similar to real simple cells, the response of the CORF model cell decreases with an increasing contrast of the mask stimulus. The linear GF model is not affected by any of the orthogonal mask stimuli

orientation preferred by the cell inhibits the response. The proposed CORF model cell also has this property as illustrated in Fig. 11. Linear summation models, such as the GF, do not possess this property, because an orthogonal mask signal contributes equally to the excitatory and inhibitory areas

Fig. 12 (*First row*) Images of natural objects from the RuG dataset. (*Second row*) The corresponding contour maps designed by a human observer. Optimal contour maps obtained by (*third row*) the GF and by (*fourth row*) the proposed CORF models. **a** Rino, **b** elephants, **c** Gnu, and **d** Hyena
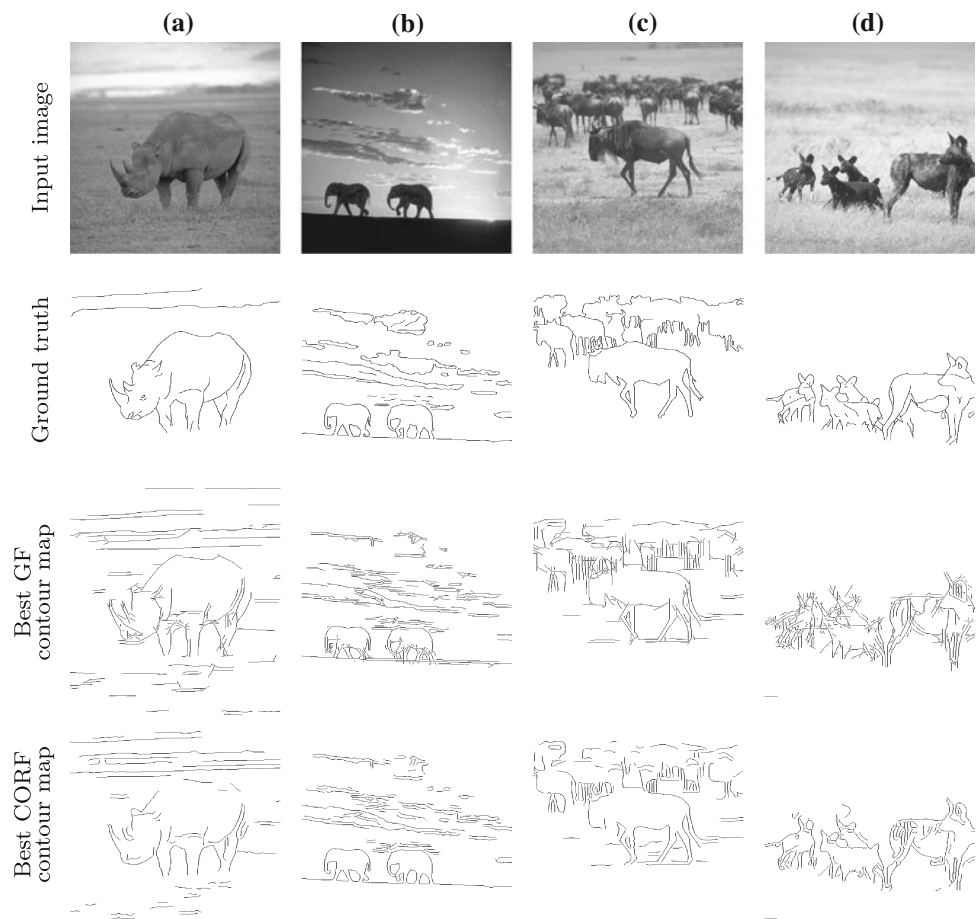
and its net contribution is zero, so that the response to the main stimulus is not affected.

These two important properties that are exhibited by the proposed CORF model are attributable to the multiplicative nature (weighted geometric mean) of the output function (Eq. 5). In Appendix B, we demonstrate that if we use an addition output function instead of multiplication, the resulting model would lose these two properties.

3.2 Performance evaluation for contour detection

In this section, we present a procedure for the evaluation of the performance of the CORF computational model in a contour detection task. We use the same procedure to evaluate the performance of the GF model and subsequently we compare the performances of the two models.

3.2.1 Data sets and ground truth

Figure 12 (first row) shows four images of natural objects and associated ground truth contour maps (second row) defined by a human observer. These images belong to a data set that

has been developed at the University of Groningen (RuG)[3]. It comprises 40 such images along with their corresponding ground truths, and was first used by Grigorescu et al. (2003). The ground truth images depict only the contours of objects, while omitting edges that are caused by texture.

Figure 13 illustrates further examples of four input images taken from a data set of Berkeley (Martin et al. 2001). This data set comprises 300 images of natural objects, and for each image the ground truth is provided as a collection of multiple (5–10) contour maps that are drawn by different human observers, Fig. 13 (second row).

In the following, we explain how we obtain contour maps from outputs of the concerned GF and CORF models. We then define performance measures that we use to evaluate a contour map obtained by a model with a corresponding ground truth contour map made by a human observer (Zhang 1996; Shin et al. 1998; Grigorescu et al. 2003).

3.2.2 Binary contour map

To obtain a binary contour map for a given input image we apply to the output of the concerned model (CORF or GF) a

---

[3] The complete data set is available at http://www.cs.rug.nl/~imaging.

Fig. 13 (*First row*) Images of natural objects from the Berkeley data set. (*Second row*) Ground truth illustrated as a superimposition of the collection of multiple contour maps drawn by different observers. *Darker contours* correspond to the agreement of multiple *hand-drawn contours*. Optimal contour maps obtained by (*third row*) the GF and by (*fourth row*) the proposed CORF models. **a** 302008, **b** 102061, **c** 302003, and **d** 368078

procedure widely used in computer vision. It consists of edge thinning by non-maxima suppression followed by binarization by hysteresis thresholding (Canny 1986; Sonka et al. 1999). The former step essentially determines the ridges in the operator response. The latter step requires two-parameter values, referred to as the low and high thresholds. We work with a value of the low threshold that is a fixed fraction (0.5) of the value of the high threshold. The resulting contour map depends on the value of the (high) threshold used: the lower that value, the larger the number of contour pixels in the map because weaker responses can pass the threshold.

The third and the fourth rows of Figs. 12 and 13 show the contour maps obtained in this way from the outputs of the GF and CORF models, respectively. These maps are obtained for certain values of the threshold parameter that are explained below.

### 3.2.3 Quantitative performance measure

We use performance indicators called recall R and precision P to measure the similarity between the contour map obtained by a given model and the ground truth. Recall R is defined as the fraction of true contour pixels (according to the ground truth) that are successfully detected by a given model. Precision P is defined as the fraction of true contour pixels from all the detected ones. They are formally defined as:

$$R \overset{\text{def}}{=} \frac{n_{TP}}{n_{TP} + n_{FN}}, \qquad P \overset{\text{def}}{=} \frac{n_{TP}}{n_{TP} + n_{FP}} \qquad (8)$$

where n_{TP}, n_{FP}, and n_{FN} stand for the numbers of true positives, false positives, and false negatives, respectively. The values of the recall and precision depend on the threshold

Fig. 14 Maximum *F*-measure values obtained with the CORF and GF models for the images of the RuG data set
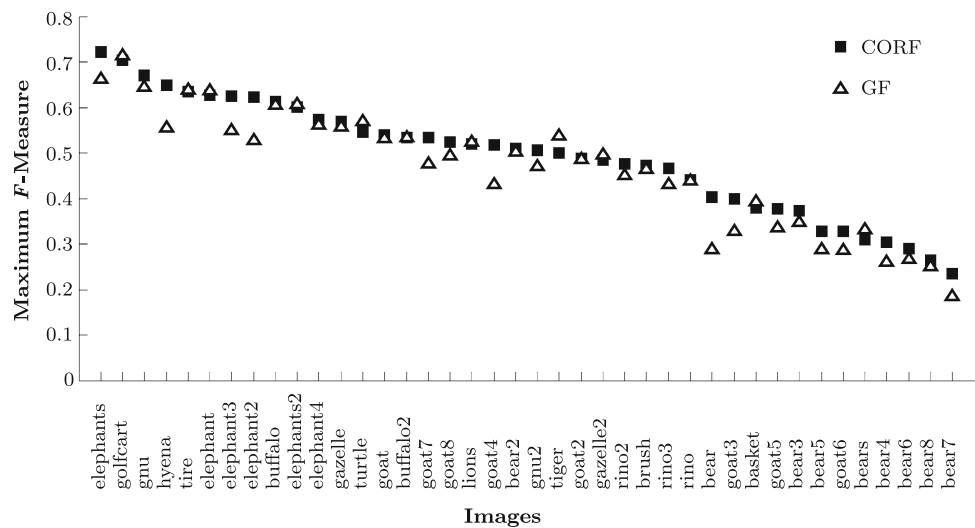
used for binarization: precision increases and recall decreases with an increasing value of that threshold.

In order to come to a single performance measure that would allow for results comparison, as suggested by Martin et al. (2004), we compute the harmonic mean of recall and precision, commonly referred to as the *F*-measure:

$$F \stackrel{\text{def}}{=} \frac{2PR}{P+R} \tag{9}$$

and consider as an optimal result the filter output for which this harmonic mean reaches its maximum for a given image.[4] The contour maps shown in Fig. 12 correspond to maximum values of the *F*-measure for the corresponding images.

3.2.4 Experimental results

In our experiments, we apply both CORF and GF computational models to every input image. Below, we specify the parameter values that were used by the concerned models and then we present a statistical comparison between their performances.

The CORF and the GF models that we compare share two parameters: the number of orientations, n_θ, which we set to 12 (intervals of $\pi/6$) and a scale parameter σ. The GF model requires two other parameters, which we set as suggested by Petkov (1995): the wavelength $\lambda = \sigma/0.4$ and the spatial aspect ratio $\gamma = 0.5$. These parameter values ensure that the RF maps of the GF and CORF models are similar.

For every input image, we only consider the two maximum *F*-measure values, one for each of the two models, across all nine scales ($\sigma \in \{1, 1.5, \ldots, 5\}$) used. For every given value

of σ, a CORF model is configured with different values[5] of the parameter ρ.

Figure 14 shows a scatter plot that illustrates a comparison between the performances of the two models for the RuG data set. The names of the images are shown on the *x*-axis in a descending order of the performance achieved with the CORF model. For 30 out of 40 input images, the maximum *F*-measure value that is achieved by the proposed CORF model is greater than the maximum *F*-measure value that is achieved by the GF model. For the images where the GF model performs better, the difference in the corresponding *F*-measure values is minimal.

We apply a right-tailed paired-samples *t* test to the set of pairs of *F*-measure values that are achieved by the proposed CORF model and the GF model and obtain as a result that the proposed CORF model outperforms the GF model for both the RuG ($t(39) = 4.39$, $p < 10^{-4}$) and the Berkeley ($t(299) = 3.88$, $p < 10^{-4}$) data sets.

Notable is the fact that, compared to the GF model, the best contour maps obtained by the proposed CORF model, such as the ones shown in the fourth rows of Figs. 12 and 13, contain less texture and the high curvature points are better preserved.

4 Discussion

The ground truth maps of the Berkeley data set were designed to evaluate performance for region segmentation rather than object contour detection. For instance, the ground truth maps shown in Fig. 13 omit the contours of the stripes in the shirt

---

[4] We refer to Appendix C for specification of the way in which we deal with inexact contour localization.

[5] For $\sigma \in \{1, 1.5, 2\}$ we use three radii ($\rho \in \{3, 7, 14\}$), for $\sigma \in \{2.5, 3, 3.5\}$ we use four radii ($\rho \in \{3, 6, 13, 25\}$) and for $\sigma \in \{4, 4.5, 5\}$ we use 5 radii ($\rho \in \{3, 5, 9, 18, 34\}$).
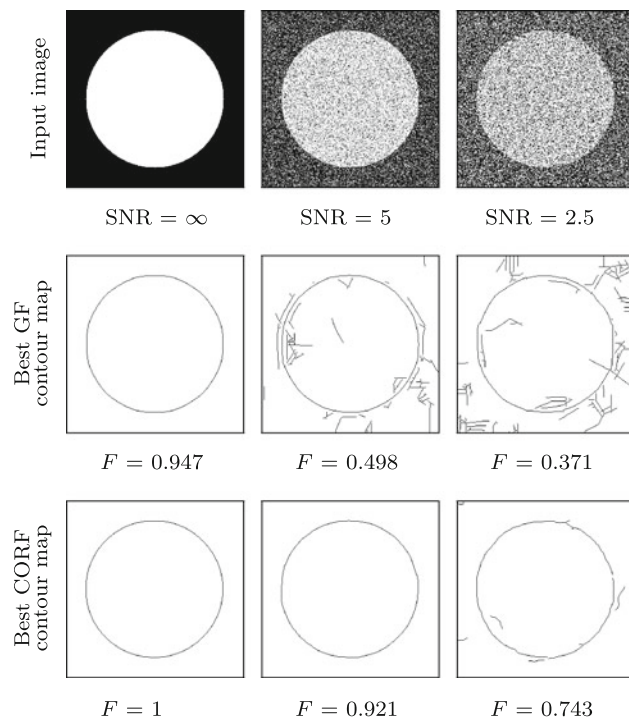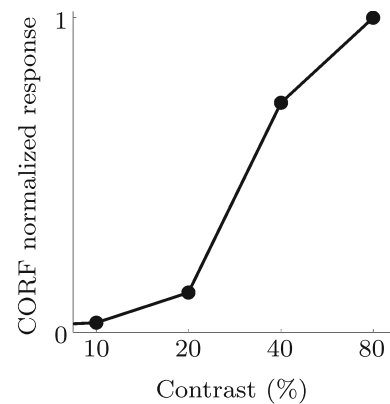
Fig. 15 (*First row*) one noiseless and two noisy stimuli with different signal-to-noise ratios (SNRs). Optimal contour maps achieved by the (*second row*) GF model and the (*third row*) proposed CORF model

Fig. 16 Responses of a CORF model cell as a function of stimulus contrast for the orientation preferred by the model. Here, the responses of the constituent model LGN cells are modified by a sigmoid function

(Fig. 13a) and the patterns in the hat (Fig. 13c) of the the two persons.

Figure 15 shows the optimal contour maps achieved by the proposed CORF model and the GF model for a synthetic stimulus without and with noise. The CORF model is more robust to noise and achieves better edge localization than the GF model. As pointed out by DuBuf (1993), curved edges that are detected by the GF model are orientation-quantized and appear as line segments in the resulting contour maps. In contrast, the proposed CORF model is able to preserve the smoothness and orientation of contours. While a high curvature contour is correctly detected by the CORF model, the GF model detects such a contour as multiple extended edges that cross each other, see, for instance, the difference in how the contours of the stripes in the shirt (Fig. 13a) and the pattern in the hat (Fig. 13c) are detected by the two models.

The class of simple cells is rather broad (DeAngelis et al. 1995). For instance, many simple cells have non-classical RFs which exhibit inhibition (Grigorescu et al. 2003). These effects are outside the scope of this study. Furthermore, the scope of the proposed CORF computational model is limited to simple cells that respond to static stimuli.

The responses of real simple cells are often not proportional to stimulus contrast. This means that the response does not double if the contrast of the stimulus grows, for instance, from 40 to 80 %. This property is known as response

saturation (Sclar et al. 1990). We can also incorporate this nonlinear property in our model by applying a sublinear function with saturation, such as the sigmoid function, to the model LGN responses. The resulting CORF model cell shows dependence of response on contrast that is qualitatively similar to the one shown by real simple cells, Fig. 16.

While we use geometric mean (essentially multiplication) as an output function of the proposed CORF computational model, we did not name it *Product* but more generally *Combination* of RF. We did that to keep the possibility open to use other output functions in the future.

The implementation of the proposed CORF model is straightforward: it includes convolutions of computed LGN responses to compute sub-unit responses,[6] shifting appropriately these sub-unit responses to take into account that they need to be taken in different points, and a pixel-wise evaluation of a multivariate output function. Elsewhere, we refer to this type of filter, built on top of simpler filters, as Combination Of Shifted Filter Responses, abbreviated as COSFIRE (Azzopardi and Petkov 2012).

The popularity of the GF model is largely due to the fact that a GF is a product of two well-known basic functions, a cosine function and a Gaussian function that we all know since secondary school. Consequently, it can be easily remembered and its properties are easily understood. At a higher level, its elegance comes from the joint maximization of its compactness in the space domain and the spatial frequency domain (Daugman 1985). It is amazing that such a conceptually simple and elegant function is a good first approximation to the properties of real neurons. The fact that it cannot reproduce all properties of real simple cells is just another evidence that the real world is less Platonic than we might wish it to be.

---

[6] This is equivalent to blurring the computed LGN responses.

5 Summary

In this paper, we proposed a CORF computational model of an orientation-selective cell in area V1 of visual cortex. Orientation selectivity is achieved by combining in an AND-type operation the responses of a collection of model LGN cells with center–surround RFs that are properly aligned.

By means of simulated reverse correlation, we demonstrated that the RF of the proposed model cell can be divided in parallel elongated excitatory and inhibitory regions, an organization that is characteristic of simple cells. We also showed that, similar to real simple cells, the response of a CORF model cell to a drifting grating alternates between high and low values with a changing phase offset of the grating. Moreover, the proposed CORF model cell exhibits important nonlinear properties, namely cross orientation suppression, contrast invariant orientation tuning and response saturation, which are found in real simple cells. Such properties are not possessed by models that are based on linear spatial summation, such as the GF model. Therefore, the proposed CORF computational model should be seen as a more realistic model of a simple cell.

We demonstrated the effectiveness of the proposed CORF model in a contour detection application and compared its performance with that of the popular GF model. For performance evaluation, we used two public data sets of natural images (RuG and Berkeley) with associated ground truth contour images. In both cases, the proposed CORF model outperforms the GF model (RuG: $t(39) = 4.39$, $p < 10^{-4}$ and Berkeley: $t(299) = 3.88$, $p < 10^{-4}$).

6 Conclusions

The proposed CORF model of a simple cell that relies on input from model LGN cells is anatomically more realistic than the GF model. Furthermore, it shares more properties with real simple cells than the GF model, such as cross orientation suppression, contrast invariant orientation tuning and response saturation. Finally, the proposed CORF model is more effective than the GF model in contour detection, which is assumed to be the primary biological task of simple cells.

Appendix A: Simulated reverse correlation

Reverse correlation (de Boer and Kuyper 1968; Jones and Palmer 1987; Ringach and Shapley 2004) is an electrophysiological technique in which a randomly generated binary or, more generally, white noise image is presented and the response of a cell to this stimulus is measured. Many different such images are presented and the images which elicit spikes from a given neuron are added together. Finally, the average intensity value is subtracted from this accumulated image. The intensity of a pixel in the resulting image is indicative for the contribution of that position in the visual field to increase (if the pixel value is positive) or decrease (if it is negative) the probability of firing of the concerned cell. Reverse correlation can be applied to determine the spatio-temporal profiles of visual neurons by taking into account the delay of a spike after a stimulus image is presented and computing a separate summation image for each delay bin. We are not concerned with temporal aspects here.

We use *simulated* reverse correlation by which we mean that we present to our CORF model cell random binary images and measure the computed response of the model cell in a given location. This response is, however, not binary (spike or no spike) as with real neurons but a graded value which is intended to represent an approximation of the firing rate of a neuron. To take this aspect into account, we weight the presented binary stimulus image with the response which the model cell produces in the considered point and sum up these response-weighted binary input images. We subtract from the response-weighted sum of binary input images its mean across all pixels. This is the way in which we obtained the 2D function shown in Fig. 8a, b.

Appendix B: Nonlinear versus linear CORF response

Figure 17 shows the orientation tuning curves of a linear CORF model in which the response is computed by summing up the responses of all afferent sub-units. This model produces response for all orientations of the presented stimulus, in contrast to the properties of real simple cells. If a threshold is used to constrain the response, the orientation bandwidth of the model becomes dependent on the contrast of the stimulus, while the orientation bandwidth of real simple cells does not depend on the contrast.

Moreover, a linear CORF model does not possess the property of cross orientation suppression. On the contrary, a mask stimulus that is orthogonal to a simultaneously presented test stimulus of the orientation preferred by the model increases rather than suppresses the response, Fig. 18.

Appendix C: Evaluation of inexact contour localization

Localization of detected contours may deviate slightly from the desired position specified in the corresponding ground truth. This might either be due to the inaccurate ground truth itself that is designed by a human observer, or otherwise by the inaccurate detection localization of the operator. For this reason, we use the tolerance method proposed by Grigorescu et al. (2004), where a detected pixel is considered correct if a corresponding ground truth contour pixel is present in a 5×5
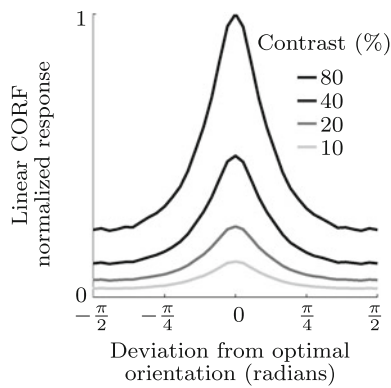
Fig. 17 Orientation tuning curves at four different contrasts obtained for a linear CORF model, the output of which is computed as the sum of the responses of all afferent sub-units. This linear CORF model gives response even at orientations that are orthogonal to the preferred orientation
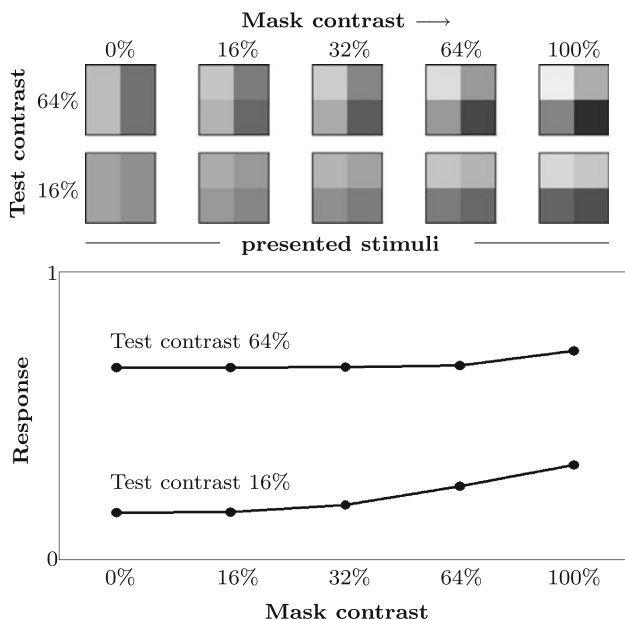
Fig. 18 A linear CORF model does not exhibit cross orientation suppression, but rather it shows the opposite effect. The plots illustrate the responses of a linear CORF model to the images shown at the *top*, compare with Fig. 11

neighborhood. Every ground truth contour pixel is considered only once, such that no two or more detected contour pixels can be matched to the same ground truth contour pixel. The false negatives and false positives can then be determined from the remaining unmatched contour pixels.

In the case of the Berkeley data set, where the ground truth is provided as a collection of multiple contour maps, we apply the following procedure. First, we match the operator contour map separately with every contour map in the corresponding ground truth collection. Subsequently, we compute the number of true positives as the total sum of the correctly matched boundary pixels with the set of ground truth contour maps.

Similarly, the number of false negatives is the total sum of incorrectly unmatched boundary pixels. However, the false positives are only those boundary pixels that do not match any of the human observer maps. For further details we refer to Martin et al. (2004).

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

Albrecht DG, De Valois RL, Thorell LG (1980) Visual cortical-neurons—are bars or gratings the optimal stimuli. Science 207(4426):88–90

Andrews BW, Pollen DA (1979) Relationship between spatial-frequency selectivity and receptive-field profile of simple cells. J Physiol 287:163–176

Azzopardi G, Petkov N (2012) Trainable COSFIRE filters for keypoint detection and pattern recognition. IEEE Trans Pattern Anal Mach Intell (to appear)

Canny J (1986) A computational approach to edge-detection. IEEE Trans Pattern Anal Mach Intell 8(6):679–698

Chung S, Ferster D (1998) Strength and orientation tuning of the thalamic input to simple cells revealed by electrically evoked cortical suppression. Neuron 20(6):1177–1189

Croner LJ, Kaplan E (1995) Receptive-fields of P-ganglion and M-ganglion cells across the primate retina. Vis Res 35(1):7–24

Daugman JG (1985) Uncertainty relation for resolution in space, spatial-frequency, and orientation optimized by two-dimensional visual cortical filters. J Opt Soc Am A 2(7):1160–1169

de Boer R, Kuyper P (1968) Triggered correlation. IEEE Trans Bio-med Eng 15(3):169–179

De Valois RL, Albrecht DG, Thorell LG (1978) Cortical cells: bar and edge detectors, or spatial frequency filters? In: Cool SJ, Smith I E L (eds) Frontiers in visual science, Springer-Verlag, Berlin, West Germany pp 544–556

De Valois KK, De Valois RL, Yund EW (1979) Responses of striate cortex cells to grating and checkerboard patterns. J Physiol 291:483–505

De Valois RL, Yund EW, Hepler N (1982) The orientation and direction selectivity of cells in macaque visual-cortex. Vis Res 22(5):531–544

DeAngelis GC, Ohzawa I, Freeman RD (1995) Receptive-field dynamics in the central visual pathways. Trends Neurosci 18(10):451–458

DuBuf JMH (1993) Responses of simple cells—events, interferences, and ambiguities. Biol Cybernet 68(4):321–333

Ferster D, Chung S, Wheat H (1996) Orientation selectivity of thalamic input to simple cells of cat visual cortex. Nature 380(6571):249–252

Finn IM, Priebe NJ, Ferster D (2007) The emergence of contrast-invariant orientation tuning in simple cells of cat visual cortex. Neuron 54(1):137–152

Gabor D (1946) Theory of communication. J Inst Electr Eng 93:429–457

Glezer VD, Tsherbach TA, Gauselman VE, Bondarko VM (1980) Linear and non-linear properties of simple and complex receptive-fields in area-17 of the cat visual-cortex—a model of the field. Biological Cybernetics 37(4):195–208

Grigorescu C, Petkov N, Westenberg MA (2004) Contour and boundary detection improved by surround suppression of texture edges. Image Vis Comput 22(8):609–622

Grigorescu C, Petkov N, Westenberg MA (2003) Contour detection based on nonclassical receptive field inhibition. IEEE Trans Image Process 12(7):729–739

Heitger F (1995) Feature detection using suppression and enhancement. Communication Technology Laboratory, Swiss Federal Institute of Technology, Lausanne, Technical Report TR-163

Hubel DH (1982) Exploration of the primary visual-cortex, 1955–78. Nature 299(5883):515–524

Hubel DH, Wiesel TN (1962) Receptive fields, binocular interaction and functional architecture in cats visual cortex. J Physiol 160(1):106–154

Hubel DH, Wiesel TN (1974) Sequence regularity and geometry of orientation columns in monkey striate cortex. J Comp Neurol 158(3):267–294

Irvin GE, Casagrande VA, Norton TT (1993) Center surround relationships of magnocellular, parvocellular, and koniocellular relay cells in primate lateral geniculate nucleus. Vis Neurosci 10(2):363–373

Jones JP, Palmer LA (1987) An evaluation of the two-dimensional gabor filter model of simple receptive-fields in cat striate cortex. J Neurophysiol 58(6):1233–1258

Kovesi P (1999) Image features from phase congruency. Videre 1(3)

Kulikowski JJ, Bishop PO (1981) Fourier-analysis and spatial representation in the visual-cortex. Experientia 37(2):160–163

Macleod IDG, Rosenfeld A (1974) Visibility of gratings—spatial frequency channels or bar-detecting units. Vis Res 14(10):909–915

Maffei L, Morrone C, Pirchio M, Sandini G (1979) Responses of visual cortical-cells to periodic and non-periodic stimuli. J Physiol 296:27–47

Marcelja S (1980) Mathematical-description of the responses of simple cortical-cells. J Opt Soc Am 70(11):1297–1300

Martin D, Fowlkes C, Tal D, Malik J (2001) A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In: Proceedings of the 8th international conference computer vision, vol 2, pp 416–423

Martin DR, Fowlkes CC, Malik J (2004) Learning to detect natural image boundaries using local brightness, color, and texture cues. IEEE Trans Pattern Anal Mach Intell 26(5):530–549

Mehrotra R, Namuduri KR, Ranganathan N (1992) Gabor filter-based edge-detection. Pattern Recogn 25(12):1479–1494

Morrone MC, Burr DC (1988) Feature detection in human-vision—a phase-dependent energy-model. Proc R Soc Lond B 235(1280):221–245

Morrone MC, Owens RA (1987) Feature detection from local energy. Pattern Recogn Lett 6(5):303–313

Movshon JA, Thompson ID, Tolhurst DJ (1978a) Receptive-field organization of complex cells in cats striate cortex. J Physiol 283:79–99

Petkov N (1995) Biologically motivated computationally intensive approaches to image pattern-recognition. Future Gener Comput Syst 11(4–5):451–465

Priebe NJ, Ferster D (2006) Mechanisms underlying cross-orientation suppression in cat visual cortex. Nat Neurosci 9(4):552–561

Reid RC, Alonso JM (1995) Specificity of monosynaptic connections from thalamus to visual-cortex. Nature 378(6554):281–284

Ringach D, Shapley R (2004) Reverse correlation in neurophysiology. Cogn Sci 28(2):147–166

Rodieck RW (1965) Quantitative analysis of cat retinal ganglion cell response to visual stimuli. Vision Res 5(11):583–601

Rosenthaler L, Heitger F, Kubler O, von der Heydt R (1992) Detection of general edges and keypoints. In: Sandini G (ed) Proceedings of the European Conference Computer Vision (ECCV92), pp 78–86

Sclar G, Freeman R (1982) Orientation selectivity in the cats striate cortex is invariant with stimulus contrast. Exp Brain Res 46(3):457–461

Sclar G, Maunsell JHR, Lennie P (1990) Coding of image-contrast in central visual pathways of the macaque moneky. Vision Res 30(1):1–10

Shin MC, Goldgof D, Bowyer KW (1998) An objective comparison methodology of edge detection algorithms using a structure from motion task. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1998 (Cat. No.98CB36231), pp 190–195

Sonka M, Hlavac V, Boyle R (1999) Image processing, analysis, and machine vision. Brooks/Cole, Pacific Grove

Stork DG, Wilson HR (1990) Do gabor functions provide appropriate descriptions of visual cortical receptive-fields. J Opt Soc Am A 7(8):1362–1373

Tyler CW (1978) Selectivity for spatial-frequency and bar width in cat visual-cortex. Vision Res 18(1):121–122

VonDer Heydt R (1987) Approaches to visual cortical function. Rev Physiol Biochem Pharmacol 108:69–150

Xu XM, Ichida JM, Allison JD, Boyd JD, Bonds AB, Casagrande V (2001) A comparison of koniocellular, magnocellular and parvocellular receptive field properties in the lateral geniculate nucleus of the owl monkey (*Aotus trivirgatus*). J Physiol 531(1):203–218

Xu XM, Bonds AB, Casagrande VA (2002) Modeling receptive-field structure of koniocellular, magnocellular, and parvocellular LGN cells in the owl monkey (*Aotus trivigatus*). Visual Neurosci 19(6):703–711

Zhang YJ (1996) A survey on evaluation methods for image segmentation. Pattern Recogn 29(8):1335–1346