Iterative Orientation Tuning of Simple Cells in V1: A Comparative Study of Two Computational Models for Contrast Detection in Images.

Marina Kolesnik, Alexander Barlit
Fraunhofer Institute for Media Communication, Schloss Birlinghoven, D-53754 Sankt-Augustin, Germany

Abstract
The orientation selectivity of simple cells in visual cortex gives a striking example of the biological system perfectly adapted to the perception of oriented stimuli. Several models have employed major principles of orientation selectivity for the processing of contrast variations in images. We have recently suggested a model for iterative orientation tuning, in which the astonishingly regular layout of simple cells is explicitly involved in the processing of oriented stimuli. In this work we extended the iterative model by incorporation a mechanism of cross-oriented inhibition. We then investigated the two models using synthetic, noisy and natural images. We found that the two models account for a large fraction of the contrast invariance of orientation selectivity – another striking aspect of the behaviour of simple cells. Our results indicate that the iterative processing of visual stimuli combined with local amplification of proximate simple cells is responsible for ~75% of the contrast invariance. Contrary to some earlier studies, the cross-oriented inhibition did not have any significant contribution to the contrast invariance but accelerated the convergence of the iterative processing on a stable solution. When probed with different images, the new model with cross-oriented inhibition generated a clear pattern of object contours.

1 Introduction

Edge detection is a cornerstone processing stage in the analysis of visual information by humans. This has triggered the development of numerous algorithms for detection of local luminance changes in images. The most popular edge detectors include Marr-Hildreth zero-crossings [1], Canny [2], Haralick [3], Deriche [4] approaches, and a full list of suggested algorithms would go on. From the sheer number of suggested edge detectors, one may conclude that the detection of contours of objects in images is not an easy task for a computer. The fact that a person can effortlessly find the contours of objects has inspired the investigation, and modelling of a contrast detection circuitry in mammalian visual system. This work attempts at developing a processing algorithm for contrast detection in images built upon principles of the physiological orientation selectivity in mammals.

Forty years ago, Hubel and Wiesel ([5], 1962) discovered that simple cells in cat primary visual cortex (V1) are tuned for the orientation of light/dark borders. The inputs to V1 come from the lateral geniculate nucleus (LGN), whose cells are not significantly orientation selective [6]. LGN cells themselves get their input from the retinal ON and OFF ganglion cells with centre-surround receptive fields (RFs), first discovered by Kuffler ([7], 1953). The orientation selectivity of simple cells in V1, as proposed by Hubel and Wiesel [5], derives from an oriented arrangement of input from the lateral geniculate nucleus (LGN): ON-centre LGN inputs have receptive fields (RFs) centres aligned over simple cell’s ON subregions, and similarly for OFF-centre inputs. Because of this input arrangement, simple cells perform a linear spatial summation of light intensity in their fields and have an elongated shape of their RFs. The orientation preference of simple cells is quite narrow, and turning a bar-stimulus through more than about 20° from the preferred orientation, greatly reduces the cell’s firing rate.

A traditional feed-forward model of the orientation selectivity performs linear spatial summation of input signals from the LGN followed by a non-linear rectification stage, in which a threshold filters out small inputs evoked by improperly oriented stimuli [8], [9]. Although many aspects of simple cell responses are consistent with this linear model, there also are important violations of linearity. For example, scaling the contrast of a stimulus would identically scale the responses of a linear cell. At high contrasts, however, the responses of simple cells show clear saturation. Such behaviour of the simple cells is referred to as contrast invariance of orientation selectivity [10].

Several neuronal models have attempted to address the nonlinearities of simple cell responses by extending the linear model to include a gain control stage [11], [12], [13]. It is suggested the response of a simple cell is governed by shunting inhibition - the divisive normalization of the cell activity due interaction with other cells. The shunting inhibition controls the gain of the transformation of the cell’s input current to output membrane potential [14]. A followed rectification stage converts the latter into a firing rate of the cell. The models with shunting inhibition predict response saturation because the divisive normalization increases with stimulus contrast. Another class of models, which also exhibit the contrast invariance rely on the amplification of LGN input by recurrent excitation.

1 To whom correspondence should be addressed: marina.kolesnik@imk.fraunhofer.de

AISB’03 Cognition in Machines & Animals, Aberystwyth, UK, April 7-11, 2003
occurring within the cortical column [29], [30], [31], [32]. This amplification is gated selectively by intracortical inhibition and thereby sharpens weak and poorly oriented LGN input. To arrive at these results the models make an assumption that LGN synapses comprise 5%-10% of the total excitatory synapses present in layer 4. However, recordings of visually evoked membrane potential changes in simple cells [33] indicate that the LGN input is responsible for generating approximately 35% to 46% responses of simple cells.

Despite remaining controversy over the details of synaptic mechanisms underlying orientation selectivity, advances in understanding of major principles of its functionality, laid the foundation for the development of computational models for contrast detection in images [15], [16], [17]. The typical architecture of such a model is built upon a simple cell circuit, which is composed of segregated ON- and OFF-data streams interacting via mechanism of opponent inhibition as suggested by Ferster [18].

None of these models, however, explicitly employed the impressive regularity in a spatial layout of simple cells called by Hubel and Wiesel [5], the functional organisation of visual cortex. Visual cortex has a distinctive striped appearance in cross-section, caused by the arrangement of cells in layers of different densities and for this reason it is also known as the striate cortex. Simple cells that respond more strongly to stimuli in one eye than in the other, and are said to show ocular dominance, are aligned into ocular dominance stripes. Moreover, when the orientation preference of cells in the ocular dominance stripes was related to their position, an astonishingly systematic organisation has emerged: the orientation preference changed linearly with position across V1 [19], [20]. After some distance where cells had shown a systematic clockwise stepping of their orientation preferences, the sequence would reverse to anticlockwise. Hubel and Wiesel therefore suggested that orientation-selective cells are organised in columns or “slabs”, in which all cells have the same preferred orientation, and that adjacent slabs represent adjacent orientation. Because, furthermore, the orientation slabs tended to be at right angles to ocular dominance stripes, their regular structure was nicknamed as Icecube layout.

Since axons and dendrites take up a significant fraction (about 60%) of the cortical volume [21], limitations on the brain size require keeping neuronal processes as short as possible. Evolution was likely to select for developmental rules that produce orientation preference maps which are sufficiently optimised in terms of length of neuronal connections. Numerical simulations relating orientation preference maps to the length of intracortical wiring have shown that the optimised layout is the Icecube if the strength of local connections is Gaussian [22]. Therefore we assume that local interactions of spatially close simple cells within the Icecube ought to be important for their functionality. A first attempt at utilizing the regularity of orientation preference maps for contrast detection in images, has been made in [23]. It is suggested that the processing of visual input undergoes several iterative cycles. The responses of simple cells at different iterations change due to local interactions of proximate cells. The model takes advantage of the regular layout of orientation preference in a very explicit way: each simple cell is sending activation into a regular net of local connections and amplifies the activity of spatially close cells. The model achieves a significant level of contrast invariance of orientation selectivity due to the iterative amplification of cells of similar orientations at retinotopically close positions.

The work here discusses an extension of the iterative model [23] by incorporating a mechanism of orthogonal suppression of spatially close simple cells of V1. The new model accounts for another aspect of the behaviour of simple cells, namely that simple cells are subject to cross-oriented inhibition; the responses to an optimally oriented stimulus can be diminished by superimposing an orthogonal stimulus that is ineffective in driving the cell when presented alone [24], [25]. We suggest that a highly systematic wiring of simple cells in neighbouring ocular dominance bands is involved in the transition of inhibiting signals to nearby cells of orthogonal orientation.

We test the performance of both models using a selected set of two-dimensional stimuli as well as noisy and natural images. Comparison of processing results reveals a very similar behavioural pattern. Both models account for a large fraction of the contrast invariance of orientation selectivity. Our results indicate that incorporation of the cross-oriented inhibition does not significantly improve its performance in terms of contrast invariance, rather stabilises the model and accelerates its convergence on equilibrium. The new model is robust to noise and, when probed with natural images, it generates a clear pattern of contours.

2 Model 1: the iterative orientation tuning of simple cells

The first model 1 has been introduced in [23] and is built upon the idea that visual perception is a continuous process of interpretation of incoming visual data. We adopt the view that the brain has no internal representation of the outside world because it is continuously available “out there” for active perception. While the eye is fixating a particular object, the low-level processing of constant visual input may be undergoing several iterative cycles. Every moment when light hits retina it would cause a different neural activity in the underlying visual circuitry, the activity which depends on a level of current neuronal excitation. Consequently, neural responses to the same visual input at different processing cycles would vary.

This iterative approach to the low-level visual processing gains a further meaning when the role of feedback connections is considered. It is logical to suggest that feedback projections, activated at subsequent iterations, would alter cell responses to the
visual input, which itself remains constant in time. It is even more likely that the regular layout of simple cells in V1 reinforces responses of simple cells at subsequent iterations. In the model, we assume that the activation of a simple cell amplifies the activity level of proximate cells, so as to tune these neighbouring cells to a local orientation pattern. After several cycles of iterative tuning, the whole system reaches equilibrium and responses of simple cells to visual input stabilise.

A neural circuit of the model for iterative orientation tuning (Fig. 1) consists of two ON- and OFF-pathways interacting via a mechanism of opponent inhibition. Visual input is processed sequentially, first by retina-LGN followed by a simple cell circuit in V1. A key feature of the model is the iterative processing of visual input, which imitates an instance when the eye is fixating a particular object and the processing of still visual input might undergo several iterative cycles. Local intracortical interaction of simple cells changes their responses to the visual input at subsequent iterations. The local interaction of simple cells is governed by their spatial layout. We adopt an Icecube model to describe the spatial layout of simple cells. In the process of local interaction the activation of each simple cell causes excitation of close cells in the Icecube layout.

The first processing stage deals with responses of retinal ganglion cells with centre-surround receptive fields (RFs) [7]. The retinal ganglion cells are modelled at each spatial position by the difference of input stimulus and its convolution with a 2-dimensional Gaussian kernel (A1), [27]. Retinal ON and OFF ganglion cells synapse mainly onto respective ON and OFF cells of the LGN. In the model, retinal inputs do not change while passing through the LGN.

Simple cells of V1 are driven by oriented input from the LGN. Physiological studies on simple cell responses recorded in cat striate cortex suggest that elongated sensitivity profile of a simple cell subfield is best modelled by a difference of two elongated Gaussians (A2). Each simple ON cell receives excitatory input from the LGN ON cells beneath it and is inhibited by LGN OFF cells at the same retinotopic position (A3).

In addition, simple cells undergo local interaction, which amplifies the activity of cells belonging to a same channel. All simple cells are considered to be stacked into a 3-D array (Icecube layout, Fig. 2), in which two coordinates define the spatial (retinotopic) position of the cell and the remaining third coordinate is related to the cell’s preferred orientation. Each simple ON cell undergoes additive amplification received from those simple ON cells that are spatially close in the 3-D array; the same ON cell is inhibited by proximate simple OFF cells (A3). The reverse arrangement holds true for simple OFF cells.

Final activation of a simple cell results from the cross-channel inhibition obtained as the steady-state solution of inhibitory shunting interaction (A4), [28].

Because the strength of amplification is an inverse function of squared distance (A6), the activation of proximate cells effectively decays within the distance of 3 units. Due to the weighting factor \( \omega = 16 \) in (A6), the effect of amplification affects only one neighbouring cell in all 8 directions within the spatial layer, and about 6 neighbouring cells in the orientation column (3 orientations both up and down the column). This local amplification enhances responses of both retinotopically proximate cells and cells tuned to similar orientations.

---

**Figure 1.** The architecture of the model neural network. The network consists of two major stages - the retina-LGN stage, followed by a simple cell circuit. Triangles at the end of lines denote the excitatory input; filled-in-black triangles denote the inhibitory input. Two dashed lines show the iterative interaction of spatially close simple cells.

**Figure 2.** A schematic diagram of the spatial layout of V1 - Icecube layout. The primary visual cortex is divided into ocular dominance columns; running perpendicular to these are orientation columns. Orientation preference within columns changes systematically so that each column represents directions from 0° to 180°. The Icecube layout of simple cells is modelled by the 3-D array consisting of spatial layers and orientation columns. Each element of the array, \((i, j, \theta)\) has two spatial coordinates, \(i\) and \(j\), for position within the layer and one orientation coordinate, \(\theta\) for position in the orientation column. It is assumed that each 3-D position contains a pair of ON and OFF cells, \(S_{on}\) and \(S_{off}\). This array layout is repeated twice for both contrast polarities \(p = 1, 2\).

The local amplification changes responses of simple cells to the same visual input over time. This is mediated by the iterative processing of visual input: amplification functions (A6) for the ON and OFF cells are fed into (A3) and the processing cycle is repeated. As the model proceeds through iterations, responses of simple cells would increase. It is however important that the model reaches equilibrium and the amplification of proximate...
cells stabilises. The corresponding balancing mechanism is provided by the cross-channel interaction (A4), which does not let responses of simple cells to grow indefinitely. However, the cross-channel interaction alone cannot fully prevent a small growth of cell responses cells in the vicinity of sharp luminance changes.

3 Model 2: the iterative tuning with cross-orientation inhibition

Incorporation of the mechanism of cross-oriented inhibition into the model 2 is based on considerable experimental evidence suggesting that stimuli at non-optimal orientations suppress the background activity of simple cells [13], [24], [25]. Clearly, the cross-oriented inhibition has the potential to suppress weak responses in the vicinity of contrastive edges, thus increasing model’s robustness to noise. Similar to the local amplification, the cross-oriented inhibition is governed by the spatial layout of simple cells in V1.

Optical imaging studies on patterns of activation across a region of monkey cortex [26] revealed a regular structure of “iso-orientation” contours radiating from points of singularity (Fig. 3). Along each of these contours the orientation preference of cell is constant, hence the name - “iso-contours”. Cells at the singularities are not orientation selective. Orientations within each bundle of iso-contours range within the interval $90^\circ$ with adjacent contours representing orientations 1.125$^\circ$ apart. A complete circle around each singularity represents a rotation from $0^\circ$ to $180^\circ$. We suggest that these iso-oriented contours are the links serving the propagation of inhibitory signals to retinotopically close cells of orthogonal orientation.

The activity of each simple cell is inhibited by four cells from the neighbouring orientation columns that are tuned to orthogonal orientation (A5). This mechanism of cross-orientation suppression affects the activity of simple cells in two ways. On one hand, the activity of a simple cell is cancelled out by the activity of retinotopically close cells of orthogonal orientation if these are strongly activated. On the other hand, the response of a strongly activated cell would only be slightly suppressed by retinotopically close cells of orthogonal orientation if these are weakly activated. This cross-oriented inhibition eliminates weak responses of simple cells while sharpening their strong responses.

4 Perception of luminance changes by the two models

We investigated the behaviour of the two models through a set of computer simulations. Each model is probed with several test stimuli, illustrating typical instances of contrast variations in images. All test stimuli are two-dimensional functions. Each model proceeds through 5 iterative cycles before edge responses of simple cells are generated. Edge response $S$ is computed by rectifying the sum of activities of ON and OFF cells minus their difference:

$$S = [S_{on} + S_{off}] - [S_{on} - S_{off}]$$  \hspace{1cm} (1)

Each plot, illustrating responses of simple cells, is a one-dimensional slice through stimuli, activity levels of simple cells at selected iterations and edge responses (1). All plots, displaying the activity level of the ON and OFF cells for the model 1 (see Fig. 4a, 5a, 10a), share a common feature: regions of strong activity levels spread onto nearby areas at later iterations. The spreading is less pronounced for the model 2 (see Fig. 4c, 5c, 10b) because the cross-oriented suppression (A5) wipes out the low-level activity at tale regions of the Gaussian-shaped response of ON and OFF cells.

Ramp transition. We conducted two series of simulations investigating the strength of responses depending on the transition range and width. On average, responses of the ON and OFF cells generated by the model 2 grow less with iterations than responses generated by the model 1 (Fig. 4 a, c; and Fig. 5 a, c). This behavioural difference occurs due to the orthogonal suppression. However, both models exhibit almost identical edge response (1) to the ramp transition regardless of its range (Fig. 4, b, d and Fig.5 b, d). Our investigation shows that the perception of ramp profile by the two models depends largely on the profile’s width and much less on the range of ramp transition. The strength of edge responses decreases approximately linearly with the increasing width of the ramp transition (Fig. 6). When the ramp width exceeds 28 pixels, responses of simple cells decay completely even though the transition range is high. This behaviour is similar for both models although the trend is quicker for the model 2: it becomes insensitive to the transition ramp wider than 22 pixels.

Figure 3. Schematic presentation of relationship between ocular dominance bands and the organisation of orientation selectivity in the visual cortex (Obermayer and Blasdel [26]). Cells along “iso-oriented contour” have the same optimal orientation. Adjacent iso-contour bundles linked at points of singularity range within two complementary sectors of $90^\circ$ each (shown in black arrows). Due to this arrangement for each cell in a given bundle there exist a counterpart cell of orthogonal orientation belonging to the adjacent bundle.
The dependency of responses on the ramp range is strongly non-linear for both models. Fig. 7 shows, that the increase of the ramp range by a factor of 9 causes about 25% growth in the strength of respective edge response. This non-linear dependency is aggregated during the iterative processing due to advantageous enhancement of initially weaker responses. This result supports a great part of the contrast invariance of orientation selectivity observed experimentally.

Further analysis of curves in Fig. 7 shows that a subtle difference in the behaviour of two models appears at later iterations. Edge responses generated by the model 2 stabilise at 6th iteration. The convergence rate is slower for the model 1. The explanation for this comes from the analysis of mechanisms stabilising the two models. There exist two such mechanisms. The first one, the cross-channel inhibition (A4), is common for both models. The second one, the cross-oriented suppression (A5) which is only present in the model 2, imposes an additional constraint on the propagation of excitation onto cells which do not receive salient oriented excitation from the visual input. Also, the cross-oriented suppression accelerates the convergence of the iterative processing of model 2 when compared to the model 1.
Figure 7. Convergence of the models on a stable solution occurs quicker for stronger responses. Whereas responses to weaker stimuli continue to grow at later iterations, strong responses do not rise any more. This feature, common for both models, is responsible for the partial contrast invariance of orientation selectivity.

Bar profile. Both models capture a large narrow variation in brightness in the form of two sharp responses (Fig. 8). The responses are associated with two sides of the bar profile, which are perceived as “edges”. The same double response pattern is induced by a narrowest possible bar profile with the width of 1 pixel. Note that the iterative amplification of edge responses by the model 2 is stronger than that one of the model 1.

Grating. Simulation of responses induced by a grating composed of four equal contrast bars produces eight sharp responses at “edge” positions, each one associated with particular bar side. However, boundary responses evoked by the two external sides of the grating undergo stronger amplification at iterations than responses to any of the internal sides of grating bars. It seems that more isolated responses tend to override nearby responses of smaller or comparable magnitude. This behavioural aspect is particularly useful for the processing of noisy images. Weak responses to spontaneous luminance variations caused by noise get eliminated after several iterations. The elimination process is especially efficient in the vicinity of strong luminance changes.
the ON cell response evoked by the adjacent luminance change in the staircase. Computation of edge responses (1) results into a sharp edge response in the middle of the step’s plateau (Fig. 10, c, d). The illusory line vanishes when the width of step’s plateau exceeds 26 pixels. The perception of illusory line might be related to the well known Chevreul illusion, in which a regular luminance staircase is perceived as not perfectly uniform along the luminance plateaus.

Figure 11. Noisy input image (left) and the cross section taken at the centre of the image (right).

Figure 12. Edge responses to the noisy image (Fig. 11) after 1st, 3rd, and 5th iteration generated by the model 1 (two upper rows) and model 2 (two bottom rows). Images of edge responses are inverted. Corresponding cross sections are taken at the centre of the images. Responses to noise weaken noticeably as the models proceed through iterative cycles. The cross-oriented inhibition introduces an additional mechanism of noise suppression for the model 2 due to suppression of the activity of cells at non-optimal orientations.

Noisy input. The processing of a synthetic image of a dark rectangle on a lighter background corrupted with 50% Gaussian noise (Fig. 11) exhibits that responses to weak contrast variations caused by noise are significantly diminished after several iterations (Fig. 12). Noise reduction is especially pronounced in the vicinity of rectangle edges, where small responses are “cancelled out” by stronger responses which spread in the process of iterative tuning.

5 Processing of natural images

Edge responses to natural images clearly illustrate a common feature exhibited by the two models, namely the enhancement of isolated weak edges at subsequent iterative cycles. The reason for this behaviour is a non-linear normalization of responses due to the cross-channel inhibition: the divisive normalization tends to boost weaker responses over the stronger ones while normalising an overall magnitude of cell responses to the range [0, 1].

Figure 13. Input image of a wound and edge responses (inverted) generated by the two models at 1st, 3rd, and 5th iteration (clockwise from top left). Weak edges disappear and strong edges are amplified as the models proceed through iterative cycles. The iterative tuning diminishes spurious responses both for skin and the wound region. Edge enhancement is more pronounced for the model 2.
6 Summary and conclusions

The motivation of this work is the development of biologically justified model for contrast detection in images, the model, which has the potential to outperform purely computational approaches to edge detection. The orientation selectivity of simple cells in V1 gives us an exciting example of a biological system, which is capable of responding to oriented visual stimuli with great efficiency. We have proposed a model for the iterative orientation tuning with cross-oriented suppression. The model is an extension of the iterative orientation tuning model introduced in [23]. We have compared the performance of the two models by probing them with synthetic stimuli, synthetic noisy images and natural images.

The two major and common for both models features are the iterative processing of visual input and the local intracortical amplification of proximate cells belonging to a same channel. The local amplification explicitly exploits morphology of simple cells in V1. The contribution of local amplification to the responses of simple cells grows with iterations. This iterative amplification greatly enhances responses of both retinotopically proximate cells and cells tuned to similar orientations. Consequently, the local amplification activates a process of selective orientation tuning enhancing responses of cells of “proper” orientations at retinotopically close positions. The cross-orientation inhibition, incorporated in the model 2, is the mechanism of selective suppression, which affects locally the activity of cells receiving stimuli with no distinguished orientation.

Three processes play a major role in the generation of the contrast invariance of orientation selectivity: 1) the iterative processing of video input, 2) the local amplification, and 3) the cross-channel inhibition of activities of simple cells. These processes are responsible for a very similar behavioural pattern exhibited by the two models. We note that the incorporation of cross-oriented suppression into the model 2 only slightly improves the model’s performance in terms of contrast invariance, which, on average, remains at a level of 75%. The orthogonal suppression does not seem to play a crucial role in the generation of contrast invariance: it does not significantly affect the magnitude of edge responses. Our investigation indicates that a major contribution to the contrast invariance comes from the local intracortical amplification of responses at subsequent iterative cycles. The cross-oriented inhibition did not account for a significant part of the contrast invariance in our simulations. This conclusion contradicts earlier suggestions that intracortical inhibition tuned to the orthogonal orientation plays a major role in the generation of cortical orientation selectivity [13], [24], [25]. However, our simulations do suggest that the cross-oriented suppression sharpens responses of simple cells. It appears that this sharpening accelerates the convergence of edge responses on a stable solution.

We conclude that the cross-oriented inhibition introduces an important stabilising factor into the process of orientation tuning, but cannot preserve the contrast invariance of orientation selectivity. Additional mechanisms may therefore be involved in the generation of the contrast invariance observed experimentally in monkeys and cats.

Although neither model includes any additional mechanism for the suppression of noise, both of them have demonstrated high robustness to noisy input. It seems that a good resistance to noise is an inherent feature of the functionality of simple cells taken over by the models for free.

One final conclusion of this study is as follows: however efficient the functionality of simple cells is, neither the iterative amplification of activities of simple cells nor the cross-oriented suppression, can provide a selective extraction of object edges. It seems that additional mechanisms such as object recognition linked with memory association feedback should play a decisive role in the selective extraction of object contours.

References

13. DeAngelis G., C., Robson, J., G., Ohzawa, I., Freeman, R.,
1. Retina-LGN. Responses of retinal ganglion ON and OFF cells, $u_{ij}^+$ and $u_{ij}^-$, are given by:

$$X_i = l_i - G_S \ast l_i;$$  \hspace{1cm} (A1)

where $\ast$ is the spatial convolution operator and $G_S$ is a centre Gaussian with standard deviation $\sigma=3$ for the model 1 and $\sigma=5$ for the model 2. The Gaussian is sampled within a filter mask of 35x35 and 45x45 pixels for the model 1 and model 2, respectively. Visual input $l$ is normalised to $[0,1]$.

2. Simple cell subfields. Simple cells are modelled for twelve discrete orientations $\theta = 0^\circ, 15^\circ, \ldots, 165^\circ$, and two opposite contrast polarities $p=1, 2$:

$$D_{\theta \sigma \omega \nu} = G_{\theta \sigma \omega \nu} - G_{\theta \sigma \omega \nu}^\ast,$$

$$G_{\theta \sigma \omega \nu} = \frac{1}{2\pi} \exp \left[ \frac{1}{2} (x - \tau)^T R^T C R (x - \tau) \right]$$  \hspace{1cm} (A2)

where $x^T = (i,j)$ denotes the position $(i,j)$, $\tau^T = (\cos \theta, \sin \theta)$ is relative offset for two Gaussian lobes from their central position $(i,j)$, and the space constants $\sigma_{\omega}=1$ and $\sigma_{\sigma}=4$ define the degree of filter’s elongation.

At each position $(i,j)$, and for each orientation, $\theta$ and polarity, $p$, the model has an even symmetric simple cell with two parallel elongated parts: an ON subfield, $R_{ij,\theta,\varphi}$, which receives excitation from LGN ON cells beneath it, $u_{ij}^+$, and is inhibited by input from the LGN OFF cells at the same position, $u_{ij}^-$; and an OFF subfield, $L_{ij,\theta,\varphi}$, for which the reverse relation to the LGN channels holds true. This physiology is embodied in the equation for the ON subfield by subtracting the half-wave rectified LGN OFF channel, $u^-$, from the rectified ON channel, $u^+$, and convolving the result with the positive lob of the oriented filter, $D_{\theta \sigma \omega \nu}$. \cite{27}.

The OFF subfield, $L_{ij,\theta,\varphi}$, is constructed similarly. In addition, each ON subfield, $R_{ij,\theta,\varphi}$, receives excitatory input, $A_{ij,\theta,\varphi}$, from all simple ON cells that are spatially close to position $(i,j)$ in the Icecube layout, and is inhibited by input $B_{ij,\theta,\varphi}$ from all close OFF cells. The reverse arrangement holds true for the computation of the activation level of each OFF subfield, $L_{ij,\theta,\varphi}$. The mutual amplification-inhibition of neighbouring cells is a time varying function which is updated iteratively.

Appendix

The processing of visual input undergoes several iterative cycles, each one containing either four (model 1) or five (model 2) subsequent stages. Stage #1, #2, #3, and #5 are common for both models. Stage #4 stands for the model 2. Below we list all processing stages noting explicitly the differences between the two models when these are present.


The above considerations give rise to the following expressions for $R^n_{i,j,\theta,p}$ and $L^n_{i,j,\theta,p}$, at iteration $n$:

$$
R^n_{i,j,\theta,p} = \left[ u^n_{i,j} + A^n_{i,j,\theta,p} - u^n_{i,j} - B^n_{i,j,\theta,p} \right] \left[ D^n_{\theta,\sigma,\tau} \right]^+ \\
L^n_{i,j,\theta,p} = \left[ u^n_{i,j} + B^n_{i,j,\theta,p} - u^n_{i,j} - A^n_{i,j,\theta,p} \right] \left[ D^n_{\theta,\sigma,\tau} \right]^+
$$

(A3)

The strength of local interaction for both models, $A_{\theta,p}$, $B_{\theta,p}$, that we call amplification functions, varies over time. The values of amplification functions at initial iteration $n=0$, are set to $A_{\theta,p} = B_{\theta,p} = 0$, for all orientations and polarities. Note, that due to the offset of the positive and negative lobes of $\tau$, subfield responses are shifted from their central positions. To compensate, both subfields, $R_{i,j,\theta,p}$ and $L_{i,j,\theta,p}$, are shifted in the opposite directions, $\tau$ and $-\tau$, respectively.

3. **Cross-channel inhibition.** The activation of simple ON cell, $S^n_{i,j,\theta}$, at iteration $n$, is obtained as the steady-state solution of inhibitory shunting interaction:

$$
S^n_{i,j,\theta} = \left[ (R^n - L^n) / (1 + R^n + L^n) \right]^{+}
$$

(A4)

Here variables occur for all positions, orientations and polarities; indexes $i$, $j$, $\theta$ and $p$ are omitted to simplify notations. Activation of simple OFF cell is obtained by interchanging $R^n$ and $L^n$.

4. **Cross-oriented suppression.**

   **Model 2:** Simple cells are engaged in the cross-orientation inhibition, so that the cells' activity at position $(i,j,\theta)$, is inhibited by four neighbouring cells in the 3-D array (see caption to Fig.2) that are tuned for the orthogonal direction:

$$
\hat{S}^{n}_{i,j,\theta} = \left[ S^n_{i+1,j,\theta} + S^n_{i-1,j,\theta} + S^n_{i,j+1,\theta} + S^n_{i,j-1,\theta} \right] / 4.
$$

(A5)

where

$$
\theta = \theta + \pi / 2 \quad \text{if} \quad 0 \leq \theta < \pi / 2 \\
\theta = \theta - \pi / 2 \quad \text{if} \quad \pi / 2 \leq \theta < \pi
$$

5. **Local amplification.** At each at position $(i,j,\theta)$ in the 3-D array, the excitatory input, $A^n_{i,j,\theta}$, from proximate ON cells, is an inverse function of squared distance:

   **Model 1:**

$$
A^n_{i,j,\theta} = \mu \sum_{m=1}^{M} \sum_{l=1}^{L} \frac{S^n_{i+m,j,l,\theta}}{D^2(l,m),(i,j,\theta)}
$$

(A6)

**Model 2:**

$$
A^n_{i,j,\theta} = \mu \sum_{m=1}^{M} \sum_{l=1}^{L} \frac{S^n_{i+m,j,l,\theta}}{D^2(l,m),(i,j,\theta)}
$$

$D^2[(l,m,k),(i,j,\theta)] = \omega (l-i)^2 + (m-j)^2 + (\theta - \theta)^2$

where $\mu$ - is a scaling factor set to: $\mu = 0.18$, and $\omega$ - is a weighting factor set to $\omega = 16$. Above computations are repeated twice for both polarities. Excitatory input $B$ to an OFF-cell is obtained by substituting $S_{off}$. 