In Rogowitz, B., editor, *Human Vision and Electronic ImagingIII*, SPIE vol. 3299, in press, 1998.

Endstopped operators based on iterated nonlinear center-surround inhibition

Erhardt Barth[&] and Christoph Zetzsche[#]

[&]NASA Ames Research Center, MS 262-2, Moffet Field, CA 94035-1000.

[#]Institut für Medizinische Psychologie, Goethestr. 31, 80336 München, Germany.

ABSTRACT

In this paper we analyze the properties of a repeated isotropic center-surround inhibition which includes simple nonlinearities like half-wave rectification and saturation. Our simulation results show that such operations, here implemented as iterated nonlinear differences and ratios of Gaussians (INDOG and INROG), lead to endstopping. The benefits of the approach are twofold. Firstly, the INDOG can be used to design simple endstopped operators, e.g., corner detectors. Secondly, the results can explain how endstopping might arise in a neural network with purely isotropic characteristics. The iteration can be implemented as cascades by feeding the output of one NDOG to a next stage of NDOG. Alternatively, the INDOG mechanism can be activated in a feedback loop. In the latter case, the resulting spatio-temporal response properties are not separable and the response becomes spatially endstopped if the input is transient. Finally, we show that ON- and OFF-type INDOG outputs can be integrated spatially to result in quasi-topological image features like open versus closed and the number of components.

Keywords: endstopping, retina, ganglion cells, lateral inhibition, feedback, corner detection, curvature, nonlinear features, topological features.

1. INTRODUCTION

Endstopping is a well known property of cortical neurons in higher vertebrates[1] and also of retinal ganglion cells in more primitive creatures like the frog[2]. However, only few attempts have been made towards a theory of endstopping and related phenomena - see Zetzsche et.al.[3] for a discussion. In previous work we have suggested a classification of local signal properties based on the intrinsic signal dimensionality[4,5,6,3]. In our view local image features are usefully classified as *i0D* (intrinsically 0D, uniform intensity), *i1D* (straight patterns, e.g, straight edges, lines, gratings), and *i2D* (not i0D or i1D, e.g., corners, line ends, junctions, etc.), and we have developed design principles for i2D-operators selective to only i2D features[4,5,6,3]. Furthermore, we have argued on information-theoretical grounds that endstopped[7] i2D-operators can provide an efficient image code, and we have proposed to analyze and design such 2D-operators in a framework based on differential geometry and filter theory[4,6,3] and also by using the Volterra-Wiener formalism[8]. Of particular interest for this paper are retinal i2D-selective cells like the "bug-detector" ganglion cells in the frog[2,9]. We here briefly outline our model to argue that i2D selectivity needs to occur at an early stage (preferable in the outer plexiform layer such that the outputs of the bipolar cells are i2D selective). Such constraints lead to the question of how i2D selectivity might arise in a rather simple isotropic retinal network.

2. SURROUND INHIBITION

Among the most basic results in sensory research are those which involve a mechanism of lateral, or surround, inhibition which seems to be a straightforward implementation of the general idea that only changes matter, and which seems to occur frequently in quite different neural architectures. In this section we will first consider a few possible and simple mathematical operations for formalizing the inhibition and then study systems in which surround inhibition is involved repeatedly.

& ebarth@vision.arc.nasa.gov; http://vision.arc.nasa.gov; phone: (650) 604 1182; fax: (650) 604 0255

chris@imp.med.uni-muenchen.de; http://www.med.uni-muenchen.de/medpsy/; phone: 49-89-5996638; fax 49-89-5996615

2.1 Nonlinear subtractive inhibition

With subtraction the value of a number a is reduced by computing $a - b$ if $b > 0$. Therefore we restrict our operations to positive numbers by formalizing the inhibition of a by b as:

$$N[N[a] - N[b]] \quad , \tag{1}$$

where $N[\ ]$ denotes half-wave rectification, i.e., $N[a] = a$ if $a > 0$ and else $N[a] = 0$. Based on considerations which are outlined in Appendix 7.1 we argue that Eq. 1 is the simplest model of subtractive inhibition. Alternatively one could argue that we employ the static nonlinearity N in addition to subtraction to enable the functionality of the iteration described below (n iterations of a linear filter function L would just filter with L^n). More complex and more realistic nonlinearities are possible, an issue which we will discuss in later sections (4.1 and 5.2).

2.2 Nonlinear divisive inhibition

The second basic mathematical operation of an inhibitory nature is division. A number a is reduced by computing a/b given $b > 1$. In the previous section we used N to create the label $x \geq 0$, and we now introduce N' to create a label $x \geq 1$:

$$N'[x] = x \text{ if } x > 1;\ N'[x] = 1 \text{ if } x \leq 1, \tag{2}$$

and the simple (see Appendix 7.1) way to divide a by b is

$$N'\left[\frac{N'[a]}{N'[b]}\right]. \tag{3}$$

2.3 Iterated nonlinear inhibition

When C. F. Gauss found what is known as the Gauss-Seidel algorithm, i.e. he noticed that the iteration of a simple rule could solve a complex system of equations, he thought it was not worth mentioning and felt embarrassed by something like a rule of thumb being more powerful than the "true" mathematics. Currently, iterative methods (e.g. related to probabilistic and physical modeling[10] and solutions of nonlinear PDEs[11]) are popular in image processing. In the context of biological vision, iterative solutions are attractive or not depending on the trade-off between hardware and time. In principle, however, we would expect the system to rather use a simple (type of) mechanism twice than a complex one only once (if possible). Several vision models involve iteration, e.g., a relaxation process[12] but the tendency is to locate such mechanisms further away from the low level.

We now investigate the input/output relationships for a simple, iterated center-surround inhibition. To do so, we take the "a" and "b" in the above Eq. 1 to be the outputs of Gaussian low-pass filters of different spatial extent and define the INDOG operator as

$$f_{i+1} = N[w_1 N[g_1 \otimes f_i] - w_2 N[g_2 \otimes f_i]];\ i = 0, ..., M;\ w_1, w_2 > 0; \tag{4}$$

where $\otimes$ denotes convolution, and g_1, g_2 are Gaussian convolution kernels with different variances $\sigma_1 < \sigma_2$ (the inhibitory filter is larger) and same function integral (same height of the corresponding filter functions, i.e., if $w_1 = w_2$ and $f(x, y)$ is constant, the difference will be zero). Initially $f_0 = I$ (image intensity). Note that in this particular case the first-stage nonlinearities applied to the convolution results are ineffective, but we keep them for formal consistency and later development - see Section 4.1.

In analogy to the INDOG and by using Eq. 3 we define the INROG operator as:

$$f_{i+1} = N'\left[w\frac{N'[g_1 \otimes f_i]}{N'[g_2 \otimes f_i]}\right],\ i = 0, ..., M. \tag{5}$$

It has been shown that images can be decomposed to and reconstructed from a ratio of Gaussians (ROG) pyramid despite the ROG being a nonlinear operation[13]. DOG and ROG operations are thus similar but for the fact that the ROG operates on image contrast. Here, we invoke these operations repeatedly with the additional nonlinearities N and N'.

3. SIMULATION RESULTS

3.1 Convergence at different ratios of filter sizes and inhibition strength

Figure 1 shows results for applying Eq. 4 to the image of a square. The results are shown for the first 7 iterations and for 77 iterations to illustrate stability and convergence. In (a) the ratio of the Gaussian filters is $\sigma_2/\sigma_1 = 2$ resulting in an octave DOG filter. In (b) and (c) the ratio has been chosen to be $\sigma_2/\sigma_1 = 7$ in accordance with current physiological and psychophysical data[14]. The weights w_1 and w_2 have been initially set to $\sqrt{(\int 1)/(\int F(G_1 - G_2))}$, where F and G are the Fourier transforms of the corresponding functions in Eq. 4, and the integrals run along spatial frequency from 0 to the maximum value; and have been subsequently corrected such that the INDOG responses after 77 iterations where comparable to the initial ones (see figure caption). In (c) results where obtained with underinhibition, i.e., $w_2/w_1 = 0.77$. Note that the convergence to the corners is slower with larger surround and lower amplitude of the inhibition. The final result, however, does not depend qualitatively on these parameters. For a given number of iterations the resulting selectivity depends on the filter parameters $\sigma_{1,2}$ and $w_{1,2}$. Note that in the context of the i0D-, i1D-, i2D- classification scheme the proposed feature hierarchy is obtained by iterating the same simple rule, i.e. i0D and i1D features are sequentially filtered out with an increasing number of iterations.

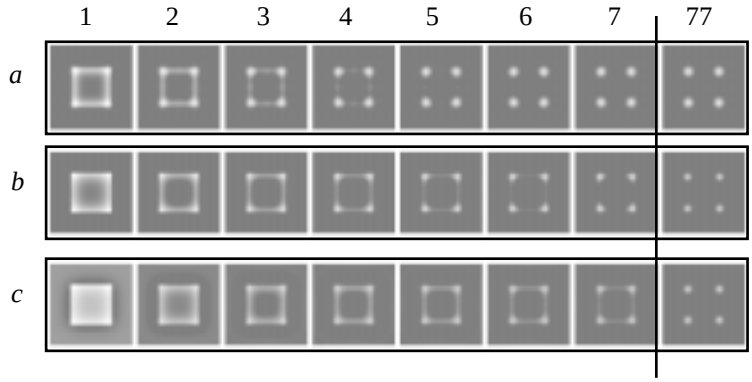


Figure 1. Results for applying Eq. 4 to a 64 by 64 pixel image of a square (top left in Figure 2) with $i = 1...7$ and 77 for the last panels. The low-pass parameters of the Gaussian filters are: $\sigma_1 = 4.87$ pixels, $\sigma_2/\sigma_1 = 2$ in (a), and $\sigma_1 = 2.04$, $\sigma_2/\sigma_1 = 7$ pixels in (b) and (c) and have been chosen such that the mid-frequencies of the resulting band-pass filters are equal in both cases. The weights are $w_1 = w_2 = 2.9$ in (a), $w_1 = w_2 = 1.3$ in (b) and $w_1 = 1.3$, $w_2/w_1 = 0.77$ in (c).

3.2 Contrast dependence of INDOG and INROG

Figure 2 shows inputs and results obtained for squares of different intensity differences and contrasts after 7 iterations of the INDOG (Eq. 4) and INROG (Eq. 5) operators. Note that the quality of the result, i.e. the 1D-/2D-classification, does not depend on contrast (compare to Figure 3 below). The response magnitudes at the corners, however, are proportional to the difference in image intensity for the INDOG and to the contrast for the INROG.

3.3 Iterations versus thresholds

Although we do not have a formal INDOG and INROG theory which would explain the convergence to i2D features, some intuition for the emerging i2D selectivity can be provided. The DOG filters response to an edge depends on the curvature of that edge. As a consequence, the response obtained for a corner and a line end (discontinuities with high curvature) is larger than for a straight edge and a straight line. The resulting difference is then enhanced by iteration. If so, why not directly apply a threshold to the first DOG such as to separate straight edges from corners. Figure 3 illustrates that the classification results obtained by thresholding depend on the image-intensity difference. Low thresholds cannot differentiate between dim corners

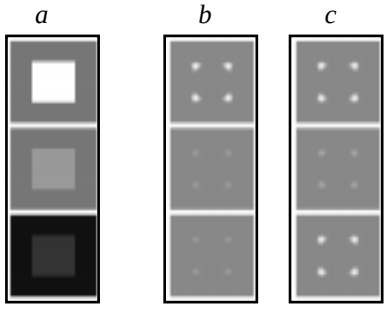


Figure 2. Input (a) INDOG with 7 iterations (b) INROG with 7 iterations (c). The grey levels in the input image are 120 and 240 (top), 120 and 150 (middle), and 30 and 60 (bottom). Thus the edges in the top and bottom image have the same contrast and those in the middle and bottom image the same difference. The NDOG parameters are as in Figure 1b. The maximum values of corner responses are (from top to bottom): 10.12, 2.53, 2.53 in (b) where grey is zero, and 1.050, 1.018, 1.050 in (c) where grey is 1.

and bright straight edges (as in Figure 3a) whereas high thresholds will only pick bright corners (as in Figure 3c). The issue of thresholding will be readdressed in Section 4.1.

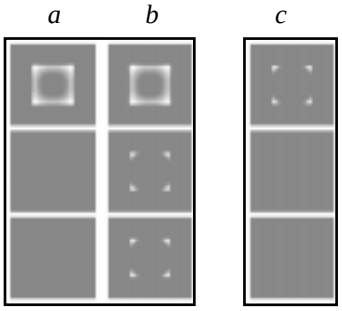


Figure 3. Thresholded NDOG responses to the same input as in Figure 2a. The values of the thresholds where 12.5 in (a) and 50 in (c) with a maximum response of 68 at the corners of the first image. For illustration purpose the panel (b) shows the same responses as (a) but with normalized output (the two lower outputs have been multiplied such that the maximum is the same for all three outputs in the column). The NDOG parameters are as in Figure 1b.

3.4 Line-length tuning

Figure 4 shows the responses of an INDOG operator to lines of different length. In addition, the line-length tuning is plotted as a function of the number of iterations. These plots exhibit the following main features: at a moderate number of iterations the response decreases with line length, with a minimum at about 3 to 4 times the optimal line width, and then increases again to a value which decreases to zero with more iterations. What this means is, that the line end itself contributes to the reduction in the response to the straight line closer to the end. This action propagates along the line and leads to a local maximum in the length tuning for a line length of about 6 to 8 times the line width. Similar plots have been obtained for a ratio of two of filter sizes and for the INROG operator.

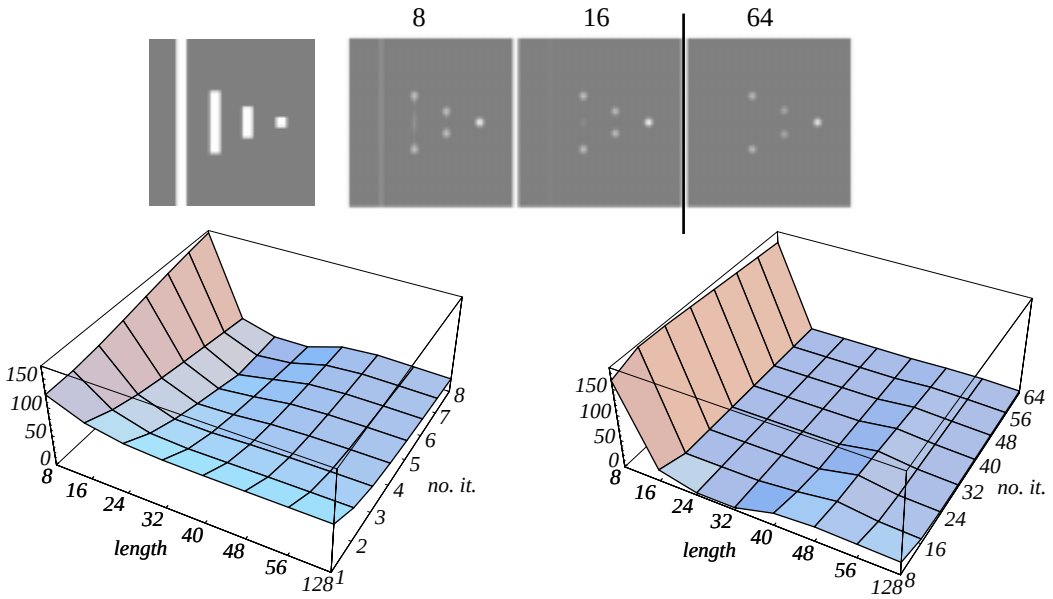


Figure 4. Line-length tuning. Input image (top left, 128 by 128 pixel) and INDOG output after 8, 16, and 64 iterations. The plots below show the line-length tuning (the maximum output values for different line length taken from the middle of the lines) for values of iterations 1 to 8 in steps of one (left) and 8 to 64 in steps of 8 (right). The width of the lines is 8 pixels and the line length varies linearly from 8 to 56 in steps of 8 pixels, and then jumps to 128 (image height). The NDOG parameters are as in Figure 1b.

3.5 Curvature tuning

Results shown in Figure 5 demonstrate that the curvature tuning of the INDOG (results for INROG are similar) is steep, with a high sensitivity to the line ends and undifferentiated response to different curvatures of the circle segments. With an increasing number of iterations the INDOG becomes a discontinuity detector selective to only blobs, corners, junctions, line ends, etc. In this property the isotropic i2D operators discussed here differ significantly from i2D operators based on oriented filters which have been used to model cortical endstopped cells and which can be designed with various curvature tunings[4,5, 6,3,15].

Figure 5. Curvature tuning. Input image (left, 256 by 128 pixel) and INDOG output after 8, 16, and 64 iterations. The NDOG parameters are as in Figure 4 (in relative units fixed by line width which here is 6 pixel). Note that similar output values are obtained for different curvatures of the circle segments, and that these values decrease with iteration.

4. QUASI-TOPOLOGICAL SELECTIVITY

We have shown before[5,9] that spatial integrals of the clipped-eigenvalues operator (an i2D operator based on the minimal eigenvalues of the Hessian) result in a quasi-topological selectivity (the selectivity is not topological because, in a strict sense, all images are topologically equivalent[9]). The essential ingredients for such a selectivity are (i) i2D-selective responses with ON and OFF separation (i.e., i2D-selective ON and OFF termed CON and COFF) and (ii) spatial pooling of CON and COFF responses such that equal amounts of CON and COFF responses cancel. We will now first show how CON and COFF operators are obtained with INDOGs and then test the quasi-topological selectivity of the mean CON-COFF response.

4.1 CON- and COFF-type INDOGs

In the previous examples we have used ON-type INDOG or INROG operators and bright-on-dark input patterns. These operators have been shown to respond "inside" a *bright* corner and line end, but they will also respond "outside" a *dark* corner or line end. This problem is always encountered in the modeling of ON/OFF type units and it is interesting to point out that the design of an ON-type operator which responds only to bright patterns is related to i2D selectivity. Consider the response of a DOG operator to an edge. If the positive part of the response is taken as ON response, an additional threshold is needed to avoid responses to dark patterns. Since at a straight edge the response is symmetric, the threshold must be higher than the response to the straight edge and will thus suppress it. In other words, an inside-outside relationship cannot be defined for a straight edge and therefore a spatial ON/OFF separation can only be obtained for curved edges. In addition, as pointed out in Figure 3, an adaptive threshold is needed to cope with the dependence on image-intensity difference and contrast of the edge and we have shown (Figure 2) that the INDOG and INROG operators can provide such an adaptive thresholding.

Guided by the intuition provided above we have found two (related) ways for obtaining an ON/OFF separation. One is to use overinhibition in Eq. 4, i.e., $w_2/w_1 > 1$. The more general approach is to introduce the generic nonlinearity $\boldsymbol{N}$, see Figure 6, and use:

$$CON: \qquad f_{i+1} = \boldsymbol{N}[\boldsymbol{N}[g_1 \otimes f_i] - \boldsymbol{N}[g_2 \otimes f_i]], \; i = 0, ..., M. \tag{6}$$

An OFF-type operator is obtained as

$$COFF: \quad f_1 = \boldsymbol{N}[\boldsymbol{N}[g_2 \otimes f_0] - \boldsymbol{N}[g_1 \otimes f_0]]; \qquad f_{i+1} = \boldsymbol{N}[\boldsymbol{N}[g_1 \otimes f_i] - \boldsymbol{N}[g_2 \otimes f_i]], \; i = 1, ..., M. \tag{7}$$

i.e., at the first iteration the nonlinearity takes the negative part of $\boldsymbol{N}[g_1 \otimes f_i] - \boldsymbol{N}[g_2 \otimes f_i]$.

$\boldsymbol{N}$, maybe differently shaped, is always present in the sense that the dynamic of the operators is limited. Linearity, as well as the nonlinearities (a) and (b) in Figure 6 are special cases of (c) with a specific setting of the working point *P* relative to the input range.

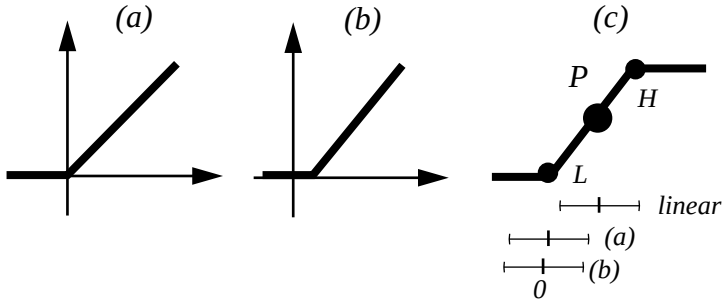


Figure 6. Different working ranges of the nonlinearity $\boldsymbol{N}$.

$CON - COFF$ responses obtained with the operators defined by Eq. 6 and Eq. 7 after only $M = 3$ iterations are shown in Figure 7 with zero as grey. The advantages of the nonlinearity of type (b) compared to (a) are the clear ON/OFF separation and a faster convergence to i2D features. Results similar to those shown in Figure 7 have been obtained by using only the second nonlinearity, i.e. by iterating $f_{i+1} = \boldsymbol{N}[g_1 \otimes f_i - g_2 \otimes f_i]$, an algorithm which might be used for technical applications since it needs the shorter programing code. In the biological context we prefer Eq. 6 because there each unit used to implement the algorithm would have the same transfer function.

4.2 Integrals over CON and COFF outputs

The numbers at the bottom of Figure 7 denote the sums $\sum_{x,y}(CON - COFF)$ for the five different input patterns in the first line divided by the sum obtained for the first image. Note that this measure is proportional to the number of "objects" and highly selective to connectedness. For more general shapes with differently curved edges the quasi-topological invariance will depend on the curvature tuning of the CON and $COFF$ operators. However, any shape can be zoomed by nearest neighbor interpolation (i.e., CON and $COFF$ computed with sub-pixel resolution) to obtain the exact invariance. In addition, contrast invariance can be obtained by using normalized CON and $COFF$ operators[5,9]. These issues are open for further investigations but it seems unlikely that a strict invariance would be desired in the context of biological vision. However, our simulations predict a high sensitivity to quasi-topological features like connectedness given that at a certain stage CON and $COFF$ units are

pooled such that equal amounts cancel out. Such a sensitivity has been found in psychophysical experiments[16,17] and also in ganglion cells of the frog - see Section 5.

Quasi-topological invariance can be obtained from pooling ON - and OFF -type units only if $\sum_{x,y}(ON - OFF) = 0$ for i0D and i1D regions[18]. In principle, it might be possible to design ON - and OFF -type operators such as to balance the sum for i0D and i1D regions. Such a strategy, however, introduces larger errors especially as it will tend to also balance the sum within i2D regions. In our example the normalized sums evaluated after the first and second iteration (straight edges are not yet suppressed but the nonlinearity is active) are: 1.00, –3.00, –2.11, 0.00, 0.76 and 1.0, –2.99, –1.30, 0.00, 0.93 (compare to the numbers in Figure 7). We conclude that spatial pooling of nonlinear ON/OFF units tends to produce a quasi-topological invariance depending on the nonlinearity. Note, however, that if the positive and negative parts of a band-pass-filter output are taken as ON and OFF the sum will be zero, i.e., equal to the DC response of the filter.

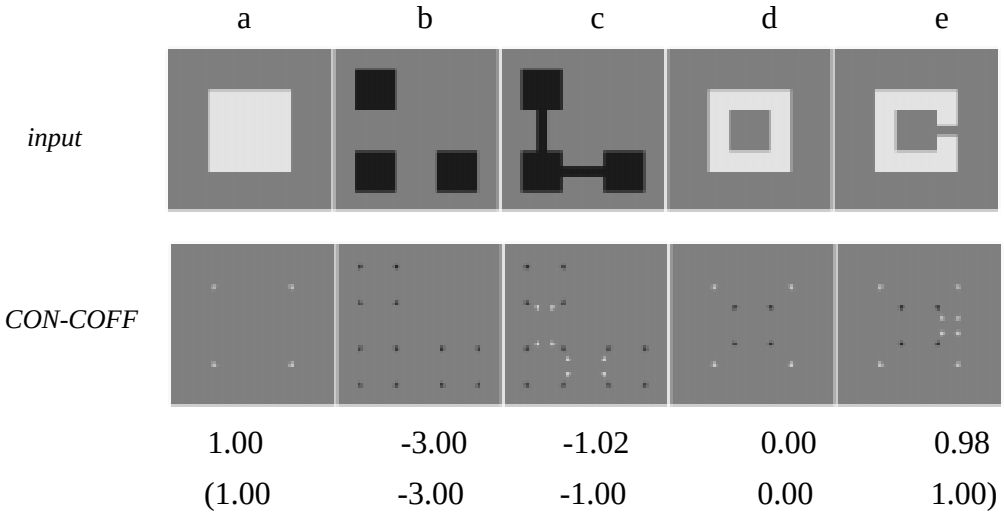


Figure 7. $CON - COFF$ responses to the input patterns in the first row. The NDOG parameters are as in Figure 1b, image size 128 by 128 pixels. The nonlinearity N is of the type shown in Figure 6b the coordinates of the points L and H in Figure 6c being $(20, 0)$ and $(1000, 1500)$ (the maximum corner response is around 200 for all 3 iterations). Numbers denote the normalized sums - see text. The numbers in brackets have been obtained with $\sigma_2/\sigma_1 = 2$.

5. RETINAL PROCESSING

Among the most remarkable ganglion cells in the frog retina are those selective to i2D features like dark blobs and corners[2,19]. The above simulations hint towards a possibility of how this can happen in a neural network with purely isotropic spatial filter characteristics. Since in the retina lateral inhibition is invoked a couple of times, a few iterations of the INDOG/INROG algorithm might be provided by cascading. But bug-detector cells are even more complex. Lettvin[20] found that the type II ganglion cells in the frog which are known to respond well to a single dark blob moved within their large receptive field do not respond to three or more such blobs. Surprisingly, the cell responds to the three blobs if the blobs are connected with lines. We were able to model this rather complex behavior of the ganglion cells by assuming a $COFF$ -type excitation with a $\sum_{x,y}(CON - COFF)$ -type inhibition[9]. For an input pattern like in Figure 7b the inhibition would be larger than for a pattern like in Figure 7c thus leading to the higher response for the connected pattern.

5.1 Cascades of high- and low-resolution NDOGS

The above model description has been implemented by using a low-resolution $NDOG$ operator acting on a high-resolution

$CON - COFF$ representation. Without going into details of the model, we argue that if the ganglion cells are to perform the low-resolution $NDOG$, then CON- and $COFF$-type selectivities must occur at earlier stages.Therefore we now consider the question of how the INDOG mechanism could be activated by feedback in the distal part of the retina (we will restrict ourselves to binary inputs and disregard the question of the inhibition being rather a difference or a ratio).

An early feedback is that from horizontal cells (HC) to cones (C). It has been first observed in 1971[21] and has subsequently "been found in all vertebrates in which satisfactory intracellular recordings from cones have been obtained" [22]. The HC/C feedback seems to mediate surround responses in second-order neurons (bipolar cells BC, ganglion cells; see Baylor et.al. [21] and Wu [24] for reviews), but the contributions of other retinal structures cannot be excluded [23,26,25]. Since it is not obvious under which conditions feedback might lead to responses similar to that of the $INDOG$, we now employ a more extensive spatio-temporal simulation of the C/HC/BC network[27].

5.2 Simulations of the distal retinal network

We have chosen to use finite-impulse response (FIR) and infinite-impulse response (IIR) filters with an additional static (sigmoid) nonlinearity (NIIR) to model the dynamical properties of the cells in the distal part of the retina, see Figure 8. Figure 9 shows the spatio-temporal model. Dotted circles indicate the spatial processing. The cones perform a spatial Gaussian low-pass filter with variance σ_e (excitatory center) while the HCs perform a Gaussian low-pass filtering with a variance which increases in time. The HC network is simulated by using low-pass filters with increasing size for each entry of the C/HC FIR filter (F4). The cones are modeled as a two-stage NIIR with the HC/C feedback at the second NIIR. HCs receive input from the cones and feed back to cones with a (constant) weighting function S_w. BCs receive excitatory input from the cones and inhibitory input from the HCs. Motivation and details of the model will be discussed in a forthcoming paper. Here we employ the model for didactical purpose to show that, in principle, i2D selectivity can occur in such a network - see Figure 10 (model parameters are specified in Appendix 7.3). Note that at stimulus offset the cone output and the BC output are i2D selective (endstopped). Depending on its temporal characteristics, a transient (temporal-OFF type) unit operating on the BC output could provide a spatial representation which is either low-pass, band-pass or i2D (endstopped). The emergence of i2D selectivity, however, depends on the nonlinearities and it is more likely to occur if the background intensity corresponds to a point P on the nonlinearity which is close to saturation (at either low or high levels).

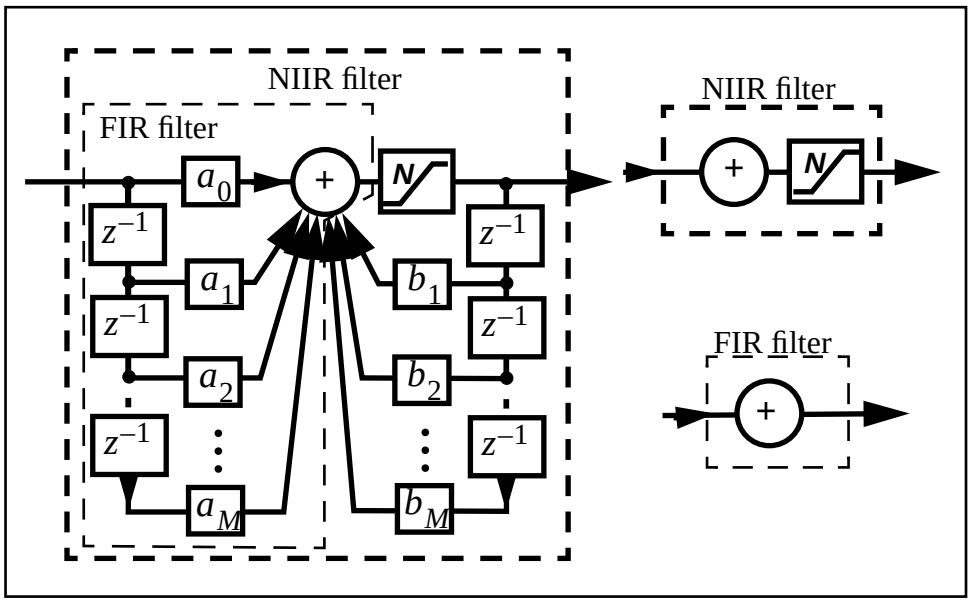


Figure 8. Model components for temporal processing. Bold and plain dashed lines define the NIIR and FIR symbols used in Figure 9.

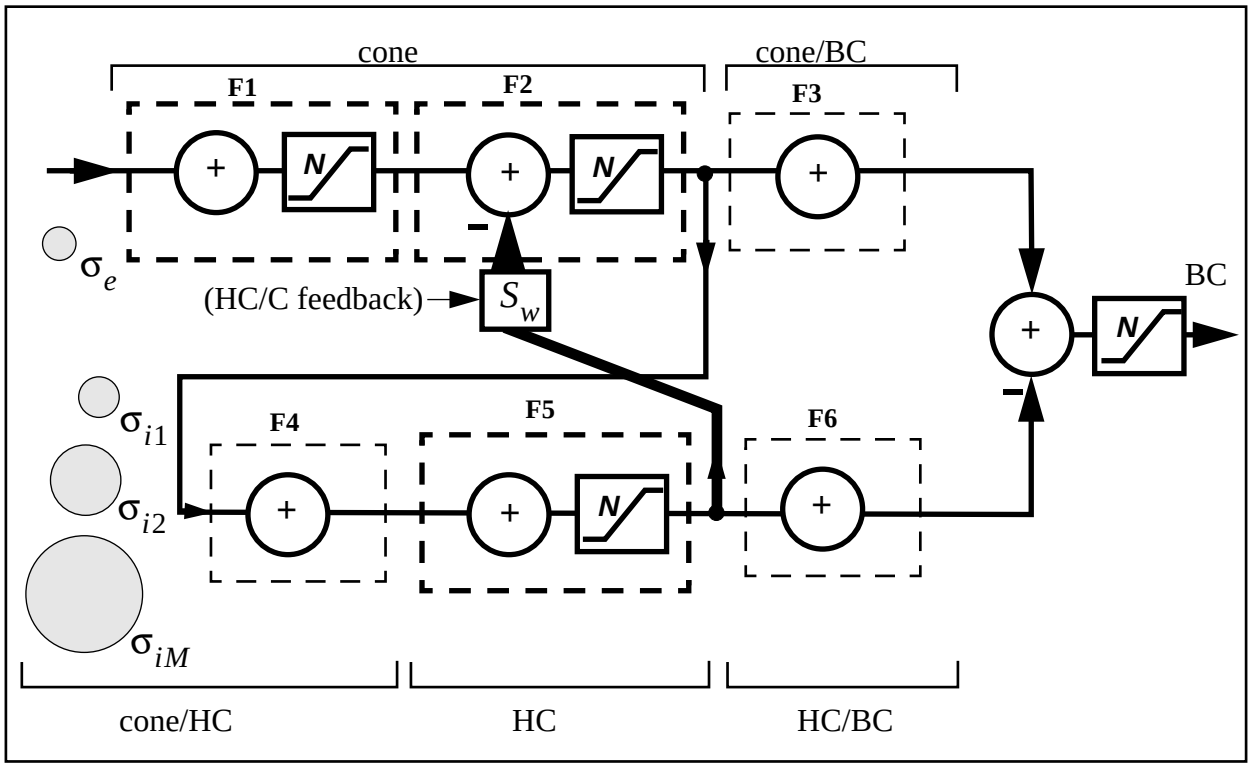


Figure 9. The spatio-temporal model for cone/HC/BC interaction.

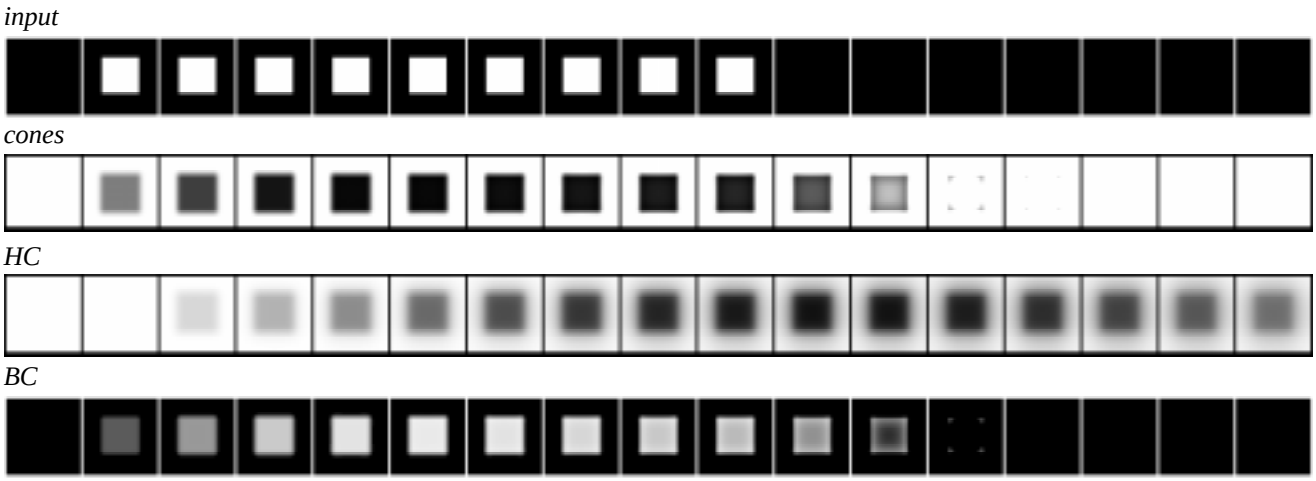


*Figure 10. Responses of model cones, HCs, and BCs to a bright square turned on and off on a dark background. The size of the image sequence is 128*128*17 pixels. See text and Appendix 7.3.*

6. DISCUSSION

The results we present in this paper have emerged from two related lines of thought in the contexts of efficient image coding and modeling of neural endstopping: (i) what are the *simplest* i2D operators, and (ii) how can i2D selectivity (endstopping) be obtained in an *isotropic* neural network.

The U-shaped line-length tuning obtained after a few iterations of the INDOG (Figure 4) is of particular interest for efficient image coding. We have argued elsewhere that, in principle, i0D- and i1D-features can be predicted from i2D features[6,3]. In practical applications, however, it might be difficult to reconstruct i1D regions between distant i2D regions. In other words, a short line will be easier to reconstruct from its line ends than a long line. Therefore a code in which short straight line- and edge segments are not coded (according to the i1D/i2D hierarchy) but longer ones are, seems a proper practical strategy.

In the context of biological vision, we conclude that i2D selectivity (and endstopping) can be obtained from neurons with isotropic centre-surround organization. The requirements are (i) iteration (cascades or feedback) and (ii) static nonlinearities. An important conclusion to be drawn from our simulations is that the probability of endstopped responses to occur increases with the number of stages involving center-surround inhibition and with the likelihood of rectifying nonlinearities (low working points on N). Furthermore, the simulations on retinal processing show that, with feedback, endstopping can occur after only one stage of inhibition. Such operations could be the basis for the i2D-selectivity of bug-detector neurons in the frog retina. For higher vertebrates, however, it seems unlikely that i2D selectivity (and endstopping) would characterize the main streams of retinal processing but it could occur under certain stimulus conditions and be enhanced by transient higher-order neurons. A certain amount of caution seems appropriate when interpreting neurophysiological data: (i) the property measured at a certain stage of the system (e.g. in V2) might have been caused somewhere else (e.g. in the retina), and (ii) the spatial-encoding properties of neurons might depend on the dynamical properties of the stimulus. However, our simulations illustrate the attractive possibility to obtain different spatial image representations at later stages by employing merely different temporal operations on the early-stage output. Finally, we have shown that pooling CON and COFF across space, creates a selectivity that is quasi topological, i.e., a cell summing such units would be selective to features like open versus closed and the number of components.

7. APPENDIX

7.1 Nonlinear simplicity

A few obvious properties of the half-wave rectification $N[\ ]$ are:

$$N[-N[-a]] = 0; \qquad N[N[a]] = N[a]; \qquad N[a] - N[b] = N[N[a] - N[b]] - N[N[b] - N[a]] \quad . \tag{8}$$

The above properties can be used to sort out further operations for computing differences between a and b:

$$N[a] - N[b] \tag{9}$$

$$|a| - |b| = N[a] + N[-a] + N[b] + N[-b] \tag{10}$$

$$a - b = N[a] - N[-a] - N[b] + N[-b] \tag{11}$$

$$N[N[a] - N[b]] - N[N[-a] - N[-b]] \tag{12}$$

$$N[N[a] - N[b]] - N[N[-a] - N[-b]] + N[N[b] - N[a]] - N[N[-b] - N[-a]] \quad . \tag{13}$$

The way these operations act on a and b is summarized in Table 1 by showing the signs and magnitudes of the resulting differences for different relationships between a and b ($a > b$, $a < b$; $a > 0$, $a < 0$; etc.).

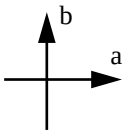
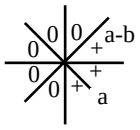
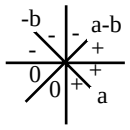
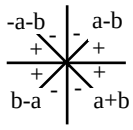
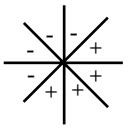
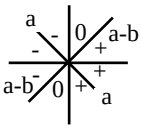
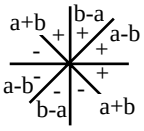
Now consider the following representation of a real number a which shows how it is constructed from positive numbers:

$$a = N[a] - N[-a] \quad . \tag{14}$$

We could argue that "a" is simpler than "$N[a]$" but also that "$N[a]$" is simpler than "$N[a] - N[-a]$" (which equals a according to Eq. 14). This illustrates that formal mathematical simplicity is easily confounded with structural simplicity with respect to a given hardware. However, all of our "$N[\ ]$" and "$N'[\ ]$" would be needless if a and b where natural numbers. At least for didactical purpose it seems useful to carry out simulations based on only natural numbers (or real numbers ≥ 1) and obtain the labels "minus" and "inverse of" by other means (e.g., different model neurons) than by extending the set of natural numbers. In this context the operations defined by Eq. 1 and Eq. 3 (for numbers $a = (N'[a]) / (N'[1/a])$ defined in

analogy to Eq. 14) are simple since they involve the smallest number of identical components.

Table 1. Resulting signs and magnitudes for different nonlinear differences.

Eq. 1	Eq. 9	Eq. 10	Eq. 11	Eq. 12	Eq. 13
						

7.2 Responses to natural images

are shown here to illustrate that the INDOG gives reasonable feature extraction.

Figure 11. $CON - COFF$ responses with $i = 1, 2, 3$ to the images shown on the left (128 by 128 pixel). Same operations and parameters as in Figure 7 were used but for the coordinates of the point L which were $(5, 0)$ (a lower threshold).

7.3 Parameters used for the example in Figure 10

- FIR and IIR filters use exponential-decay kernels with parameters: kernel size (s), time constant in frames (t), and kernel sum (n) given as (s, t, n). **F1**: (3, 2, 1), (6, 4, -0.7); **F2**: (3, 2, 1.0); (6, 4, 0.9); **F3**: (3, 2, 1); **F4**: (5, 3, 1); **F5**: (7, 4, 1.0), (9, 6, 0.2); **F6**: (3, 2, 0.3).
- Spatial filters: $\sigma_e = 6.6$ pixel; $\sigma_{i1} = 20$ pixel, doubled at each time step of **F4**.
- The nonlinearity ***N*** is sigmoid defined as $y = y_{mid} + \frac{Y_{max} - Y_{min}}{2} \cdot \tanh\left(\pi \cdot s \cdot \frac{x - X_{mid}}{X_{max} - X_{min}}\right)$, with steepness s, and mid input value $X_{mid} = \frac{X_{max} - X_{min}}{2} + X_{min}$. The settings are $X_{min} = Y_{min} = 0$, $X_{max} = Y_{max} = 255$, $X_{mid} = 128$, and $s = 1.0$ for all ***N***, but a $s = 1.4$ for the last nonlinearity at the BC output. The intensity values in the input sequence are 0 and 255.
- The feedback weight is $S_w = 0.6$.

8. ACKNOWLEDGEMENT

The work has been supported by the Deutsche Forschungsgemeinschaft grant Ba 1176/4-1 to E.B. Part of the results have been presented before[9,27] and supported by DFG grant Re 337/7 to I. Rentschler and C.Z. We thank D. Grosof, J. Mulligan, and B. Watson for comments on the manuscript.

9. REFERENCES

1. Hubel, D. H. and Wiesel, T. N. “Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat”, *J of Neurophysiology*, 28:229–289, 1965.
2. Lettvin, J. Y., Maturana, H. R., McCulloch, W. S., and Pitts, W. H. “What the frog’s eye tells the frog’s brain” *Proceedings IRE*, 47:1940–1951, 1959.
3. Zetzsche, C., Barth, E., and Wegmann, B. “The importance of intrinsically two-dimensional image features in biological vision and picture coding”, in Watson, A., editor, *Digital images and human vision*, pages 109–138. MIT Press, Cambridge, MA, 1993.
4. Zetzsche, C. and Barth, E. “Fundamental limits of linear filters in the visual processing of two-dimensional signals”, *Vision Research*, 30:1111–1117, 1990.
5. Zetzsche, C. and Barth, E. “Image surface predicates and the neural encoding of two-dimensional signal variation”, in Rogowitz, B., editor, *Human Vision and Electronic Imaging: Models, Methods, and Applications*, volume SPIE 1249, pages 160–177, 1990.
6. Barth, E., Caelli, T., and Zetzsche, C. “Image encoding, labelling and reconstruction from differential geometry”, *CVGIP:GRAPHICAL MODELS AND IMAGE PROCESSING*, 55(6):428–446, 1993.
7. “Endstopping” refers to only one aspect of i2D selectivity.
8. Krieger, G. and Zetzsche, C. “Nonlinear image operators for the evaluation of local intrinsic dimensionality”, *SPECIAL ISSUE ON NONLINEAR IMAGE PROCESSING - IEEE Trans. Image Processing*, 5(6):1026–1042, 1996.
9. Barth, E., Ferraro, M., Zetzsche, C., and Rentschler, I. “Computational models for the topological selectivity in early and primitive vision systems”, *OSA Annual Meeting Technical Digest*, 16:186, 1993.
10. Blake, A, Yuille, A. editors. *Active Vision*., MIT Press, 1982.
11. Romeny ter Haar, B. M. editor. *Geometric--Driven Diffusion in Computer Vision*., Kluwer, 1994.
12. Caelli, T. M. “An adaptive computational model for texture segmentation”, *IEEE Trans. SMC*, 18(1):9–17, 1988.
13. Zetzsche, C. and Hauske, G. “Multiple channel model for the prediction of subjective image quality”, in Rogowitz, B., editor, *Human Vision, Visual Processing and Visual Display*, SPIE volume 1077, pages 209-216, 1989.
14. Rohaly, A.M., Ahumada A.J., Watson, A.B. “Object detection in natural backgrounds predicted by discrimination performance and models”, Vision Research 37: 32225-3235, 1997.
15. Note by comparing the results in Figure 4 and Figure 5 with those in previous sections, that it takes more iterations to suppress a straight line compared to a straight edge, i.e. the speed of convergence towards i2D features depends on the resolution (and presumably the shape) of the filters; an issue which needs further investigation.
16. Chen, L. “Topological structure in visual perception”, *Science*, 218(12):699–700, 1982.
17. Elder, J. and Zucker, S. “The effect of contour closure on the rapid discrimination of two-dimensional shapes. *Vision Research*, 33(7): 981-991, 1993.
18. Consider a rectangle which is stretched along its edges without changing its topology. If the $CON - COFF$ response differs from zero at the straight edges the sum will change with stretching. As shown by the Gauss-Bonnet theorem of differential geometry, the integral curvature of a surface will change only if the topology changes.
19. Lettvin, J. Y., Maturana, H. R., Pitts, W. H., and McCulloch, W. S. “Two remarks on the visual system of the frog”, in Rosenblith, W. A., editor, *Sensory Communications*, pages 757–776. John Wiley & Sons, New York, 1961.
20. Lettvin, J. “Observed plasticities which are difficult to account”, presented at the *Symposium on Dynamic Properties of Receptive Fields and Plasticity of Processing Systems*, Triest, Italy, 1995.
21. Baylor, D. A., Fuortes, M. G. F., and O’Bryan, P. M. “Receptive fields of cones in the retina of the turtle”, *J Physiol*, 214:265–294, 1971.
22. Burkhardt, D. A. “Synaptic feedback, depolarization, and color opponency in cone photoreceptors”, *Vis Neurosci*, 10(6):981–989, 1993.
23. Kaplan, E., Lee, B. B., and Shapley, R. M. “New views on primate retinal function”, in Osborne, N. and Chader, J., editors, *Progress in retinal research*, volume 9, chapter 7, pages 273–335. Pergamon Press, Oxford, 1990.
24. Wu, S. M. “Feedback connections and operation of the outer plexiform layer of the retina”, *Curr Opin Neurobiol*, 2(4):462–468, 1992.
25. Schiller, P. H. “The ON and OFF channels of the visual system”, *Trends in Neurosciences*, 15(3):86–92, 1992.
26. Wässle, H. and Boycott, B. “Functional architecture of the mammalian retina”, *Physiological reviews*, 71(2):447–480, 1991.
27. Barth, E., Krieger, G., Rentschler, I., and Treutwein, B. “Receptor-horizontal cell interactions may induce endstopping”, in *Invest. Ophthalmology and Vis. Science*, Suppl., page 1056, 1996.