Contour Detection Operators Based on Surround Inhibition

Grigorescu, Cosmin; Petkov, Nicolai; Westenberg, Michel A.

Published in:
2003 INTERNATIONAL CONFERENCE ON IMAGE PROCESSING, VOL 3, PROCEEDINGS

Publication date:
2003

CONTOUR DETECTION OPERATORS BASED ON SURROUND INHIBITION

Cosmin Grigorescu, Nicolai Petkov, and Michel A. Westenberg

Inst. of Mathematics and Computing Science, University of Groningen
P.O. Box 800, 9700 AV Groningen, The Netherlands
E-mail: {cosmin, petkov, michel}@cs.rug.nl

ABSTRACT

We propose a biologically motivated computational step, called non-classical receptive field (non-CRF) inhibition, to improve contour detection in images of natural scenes. We augment a Gabor energy operator with non-CRF inhibition. The resulting contour operator responds strongly to isolated lines, edges, and contours but exhibits a weaker or no response to edges that make part of texture. As such, the contour operator is more useful for contour-based object recognition tasks than traditional edge detectors, which do not make such a distinction. The contour operator consistently outperforms the Canny edge detector on natural images with associated ground truth contour maps.

1. INTRODUCTION

One of the problems with contemporary edge detectors is that they do not make a distinction between isolated edges and edges originating from a textured region. There is evidence that the human visual system makes such a difference in its early stages of visual information processing, and that isolated edges, on one hand, and edges in a group, on the other hand, are perceived in different ways.

Psychophysical experiments have shown that the perception of an oriented stimulus, such as a line or a contour, can be influenced by the presence of other such stimuli (distractors) in its neighborhood, see Fig. 1(a)-(c). There is neurophysiological evidence that the response of orientation-selective neurons in the visual cortex to an optimal stimulus (a line or an edge) is reduced when other oriented stimuli are added to the surrounding [1]. This effect, called non-classical receptive field (non-CRF) inhibition, is shown by nearly 80% of all orientation selective cells. In approximately one third of the cells, the suppression effect depends on the orientation of the surrounding stimuli, whereas in 40% of the cells the effect is independent of the orientation of the surrounding stimuli [1].

In this study, we examine more closely what the biological utility of such a mechanism would be. Our main hypothesis is that this inhibitory mechanism suppresses edges which make part of texture, while it does not suppress edges that belong to the contours of objects. We propose two contour operators based on Gabor energy augmented with non-CRF inhibition, and compare their performance with the standard Canny edge detector [2].

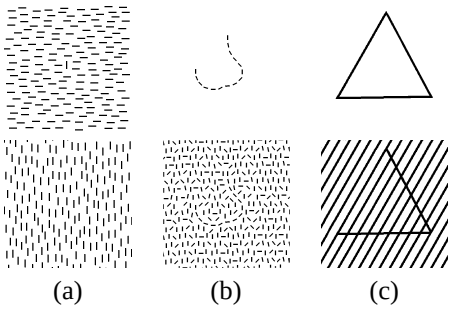


Fig. 1. Manifestations of perception modulation by the context. (a) Orientation contrast pop-out: the center stimulus segregates when surrounded by elements of orthogonal orientation, but it does not "pop out" in a surrounding of parallel stimuli. (b) An isolated contour is more salient than a contour which is surrounded by texture. (c) The perception of the three legs of the triangle depends on the context in which the triangle is embedded.

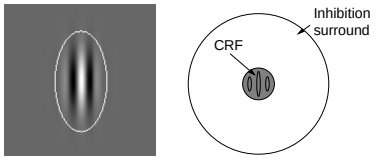
2. COMPUTATIONAL MODEL

We briefly summarize the model of non-CRF inhibition introduced elsewhere [3]. We consider two types of inhibition: (i) anisotropic, in which only responses obtained for the same preferred orientation as a central response contribute to the suppression, and (ii) isotropic, in which all responses neurons contribute to the suppression in an equal way, independently of their preferred orientations.

2.1. Simple and Complex Cells

The spatial summation properties of simple cells can be modelled by a family of 2-D Gabor functions [4]. A receptive field function $g_{\lambda,\sigma,\theta,\varphi}(x,y)$, $(x,y) \in \Omega \subset \mathbb{R}^2$, centered

0-7803-7750-8/03/$17.00 ©2003 IEEE.



(a) (b)

Fig. 2. (a) Intensity map of a 2-D Gabor function. The bright ellipse specifies the boundary of the CRF. (b) Non-CRF inhibition is caused by the surround of the CRF.

in the origin specifies the response to an impulse at point (x, y) and is defined as follows:

$$\begin{aligned} g_{\lambda,\sigma,\theta,\varphi}(x,y) &= e^{-\frac{\tilde{x}^2+(\gamma\tilde{y})^2}{2\sigma^2}}\cos(2\pi\frac{\tilde{x}}{\lambda}+\varphi) \\ \tilde{x} = x\cos\theta + y\sin\theta, &\quad \tilde{y} = -x\sin\theta + y\cos\theta, \end{aligned} \tag{1}$$

where $\gamma = 0.5$ is the spatial aspect ratio, σ determines the size of the receptive field, and λ is the preferred wavelength. The ratio σ/λ determines the spatial frequency bandwidth and the number of parallel excitatory and inhibitory stripe zones in the receptive field (Fig. 2(a)). In this paper, we use $\sigma/\lambda = 0.56$, which corresponds to a bandwidth of one octave at half-response. The angle parameter θ, $\theta \in [0, \pi)$, determines the preferred orientation. The parameter φ, $\varphi \in (-\pi, \pi]$, is a phase offset that determines the symmetry of $g_{\lambda,\sigma,\theta,\varphi}(x, y)$ with respect to the origin. The response $r_{\lambda,\sigma,\theta,\varphi}(x, y)$ of a simple cell with a receptive field function $g_{\lambda,\sigma,\theta,\varphi}(x, y)$ to an input image with luminance distribution $f(x, y)$ is computed by convolution:

$$r_{\lambda,\sigma,\theta,\varphi}(x,y) = (f * g_{\lambda,\sigma,\theta,\varphi})(x,y). \tag{2}$$

The Gabor energy $E_{\lambda,\sigma,\theta}(x, y)$ is related to a model of a complex cell which combines the responses $r_{\lambda,\sigma,\theta,0}(x, y)$ and $r_{\lambda,\sigma,\theta,-\frac{\pi}{2}}(x, y)$ of a pair of a symmetric and an anti-symmetric filter as follows:

$$E_{\lambda,\sigma,\theta}(x,y) = \sqrt{r^2_{\lambda,\sigma,\theta,0}(x,y) + r^2_{\lambda,\sigma,\theta,-\frac{\pi}{2}}(x,y)}. \tag{3}$$

In the following, we use the Gabor energies $E_{\lambda,\sigma,\theta_i}(x, y)$ for a number of N_θ different orientations

$$\theta_i = \frac{(i-1)\pi}{N_\theta}, \qquad i = 1, 2, \ldots, N_\theta. \tag{4}$$

2.2. Models of Non-CRF Inhibition

For a given point in the image, an inhibition term is computed by weighted summation of the responses in a ring-formed area surrounding the CRF centered at the concerned point (cf. Fig. 2(b)). We use a normalized weighting function $\mathrm{w}_\sigma(x, y)$ defined as follows:

$$\begin{aligned} \mathrm{w}_\sigma(x,y) &= \frac{1}{||H(\mathrm{DoG}_\sigma)||_1} H(\mathrm{DoG}_\sigma(x,y)), \\ H(z) &= \begin{cases} 0 & z < 0 \\ z & z \geq 0, \end{cases} \end{aligned} \tag{5}$$

where $||.||_1$ denotes the L_1 norm, and $\mathrm{DoG}_\sigma(x, y)$ is the following difference of Gaussians:

$$\mathrm{DoG}_\sigma(x,y) = \frac{1}{2\pi(4\sigma)^2} e^{-\frac{x^2+y^2}{2(4\sigma)^2}} - \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}. \tag{6}$$

2.2.1. Anisotropic inhibition.

We compute a suppression term $t^A_{\lambda,\sigma,\theta_i}(x, y)$ for each orientation θ_i as follows:

$$t^A_{\lambda,\sigma,\theta_i}(x,y) = (E_{\lambda,\sigma,\theta_i} * \mathrm{w}_\sigma)(x,y). \tag{7}$$

We now introduce a new operator $\tilde{b}^{A,\alpha}_{\lambda,\sigma,\theta_i}(x, y)$ which takes as its inputs the Gabor energy $E_{\lambda,\sigma,\theta_i}(x, y)$ and the inhibition term $t^A_{\lambda,\sigma,\theta_i}(x, y)$:

$$\tilde{b}^{A,\alpha}_{\lambda,\sigma,\theta_i}(x,y) = H(E_{\lambda,\sigma,\theta_i}(x,y) - \alpha t^A_{\lambda,\sigma,\theta_i}(x,y)). \tag{8}$$

The factor α controls the strength of the suppression exercised by the surround on the Gabor energy operator. Finally, we introduce the contour detector with anisotropic inhibition $b^{A,\alpha}_{\lambda,\sigma}(x, y)$, which takes the maximum response of $\tilde{b}^{A,\alpha}_{\lambda,\sigma,\theta_i}(x, y)$ over all orientations:

$$b^{A,\alpha}_{\lambda,\sigma}(x,y) = \max\{\tilde{b}^{A,\alpha}_{\lambda,\sigma,\theta_i}(x,y) \mid i = 1, \ldots, N_\theta\}, \tag{9}$$

We also compute an orientation map $\Theta^A(x, y)$ containing the orientation for which this maximum response is achieved:

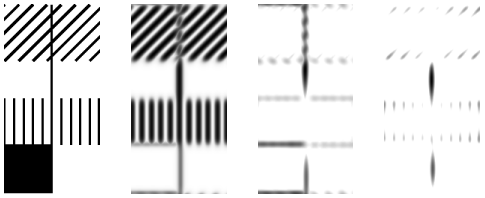
$$\begin{aligned} \Theta^A(x,y) &= \theta_k, \\ k &= \mathrm{argmax}\{\tilde{b}^{A,\alpha}_{\lambda,\sigma,\theta_i}(x,y) \mid i = 1, \ldots, N_\theta\}. \end{aligned} \tag{10}$$

Fig. 3(c) shows the output of the contour detector on a synthetic input image. Since the suppression effect depends on the relative orientation of center and surround, we refer to this type of modulation as *anisotropic inhibition*.

2.2.2. Isotropic inhibition.

We model isotropic inhibition by computing an inhibition term $t^I_{\lambda,\sigma}(x, y)$ that is independent of orientation. First, we construct an energy map $\widehat{E}_{\lambda,\sigma}(x, y)$ with values of maximum Gabor energy response:

$$\widehat{E}_{\lambda,\sigma}(x,y) = \max\{E_{\lambda,\sigma,\theta_i}(x,y) \mid i = 1, \ldots, N_\theta\}, \tag{11}$$



(a) (b) (c) (d)

Fig. 3. (a) Synthetic input image. (b) The Gabor energy operator (according to (11)) responds to lines and edges independently of the context, i.e., the surrounding in which these lines and edges are embedded. (c) The operator with anisotropic inhibition responds selectively to isolated lines and edges, and lines that are surrounded by a grating of a different orientation. (d) The operator with isotropic inhibition responds selectively to isolated lines and edges only.

and an orientation map $\Theta^I(x,y)$ with the orientation for which this maximum response is achieved:

$$\begin{aligned} \Theta^I(x,y) &= \theta_k, \\ k &= \operatorname{argmax}\{E_{\lambda,\sigma,\theta_i}(x,y) \mid i = 1,\ldots,N_\theta\}. \end{aligned} \tag{12}$$

The inhibition term $t^I_{\lambda,\sigma}(x,y)$ is computed as follows:

$$t^I_{\lambda,\sigma}(x,y) = (\widehat{E}_{\lambda,\sigma} * \mathrm{w}_\sigma)(x,y). \tag{13}$$

Finally, we introduce the contour detector with isotropic inhibition $b^{I,\alpha}_{\lambda,\sigma}(x,y)$ which takes as its inputs the maximum energy map $\widehat{E}_{\lambda,\sigma}(x,y)$ and the inhibition term $t^I_{\lambda,\sigma}(x,y)$:

$$b^{I,\alpha}_{\lambda,\sigma}(x,y) = H(\widehat{E}_{\lambda,\sigma}(x,y) - \alpha t^I_{\lambda,\sigma}(x,y)). \tag{14}$$

The factor α controls the strength of the inhibition of the surround on the maximum Gabor energy term. Fig. 3(d) shows the output of this operator on a synthetic input image.

2.3. Binary Edge Map Construction

Binary edge maps are constructed by the standard procedure of nonmaxima suppression followed by hysteresis thresholding [2]. From the orientation map $\Theta^I(x,y)$ (or $\Theta^A(x,y)$) and corresponding response $b^{I,\alpha}_{\lambda,\sigma}(x,y)$ (or $b^{A,\alpha}_{\lambda,\sigma}(x,y)$) that specify the normal to the local edge direction and the local edge strength, respectively, nonmaxima suppression thins the edges to one-pixel wide candidate contours. The final binary contour map is computed from the candidates by hysteresis thresholding. This process involves two threshold values t_l and t_h, $t_l < t_h$. Commonly, t_h is computed as a quantile $t_h(1-p)$ of the percentage p of the candidate pixels that should be retained in the final edge map. We fix the low threshold value t_l to $t_l = 0.5t_h$.

We decided for the same post-processing operations as performed by the Canny edge detector [2] in order to simplify comparison in a later stage.

3. PERFORMANCE EVALUATION

Most state-of-the-art methods that evaluate the performance of an edge detector use natural images with an associated ground truth specified by a human [5]. We selected a set of 40 images which depict either man-made objects on textured background or animals in their natural habitat; for each image, an associated ground truth binary contour map was drawn by hand. A pixel is included in the ground truth if any of the following criteria holds: (i) it is part of the outline of an object or it belongs to a contour in the interior of an object; (ii) it makes part of boundaries between (textured) regions (e.g. sky and grass, water and sky).

3.1. Performance Measure

We describe here briefly the performance measure introduced in [6]. Let E_{GT} and B_{GT} be the set of contour pixels and background pixels of the ground truth contour image, respectively, and E_{D} and B_{D} be the set of contour pixels and background pixels of the operator-detected contour image, respectively. The set of correctly detected contour pixels is $E = E_{\mathrm{D}} \cap E_{\mathrm{GT}}$. False negatives, i.e. ground-truth contours missed by the contour detector, are given by the set $E_{\mathrm{FN}} = E_{\mathrm{GT}} \cap B_{\mathrm{D}}$, while false positives (spurious contours) are given by the set $E_{\mathrm{FP}} = E_{\mathrm{D}} \cap B_{\mathrm{GT}}$. We define a *performance measure* of a contour detector as

$$P = \frac{\mathrm{card}(E)}{\mathrm{card}(E) + \mathrm{card}(E_{\mathrm{FP}}) + \mathrm{card}(E_{\mathrm{FN}})}, \tag{15}$$

in which $\mathrm{card}(X)$ denotes the number of elements of set X. Since contours cannot always be detected at exact integer image coordinates, we consider that a contour pixel is correctly detected if a corresponding ground truth contour pixel is present in a 5×5 square neighborhood centered at the respective pixel coordinates.

3.2. Experimental Results

We compare the performances of the two contour detectors with the performance of the Canny edge detector.

The Canny edge detection operator [2] computes a gradient magnitude and direction for each pixel of the filtered image, and constructs a binary edge map by post-processing with nonmaxima suppression and hysteresis thresholding. The gradient magnitude and direction are computed using a scale-dependent differential geometry operator.

The parameters of the Canny edge detector are σ, the standard deviation of a Gaussian smoothing kernel, and p, the percentage of edge pixels that should be retained in the final edge map. The contour detectors have an additional parameter, α, which is the texture inhibition factor. The parameter σ here denotes the standard deviation of the Gaussian envelope of the Gabor function. For the Canny edge

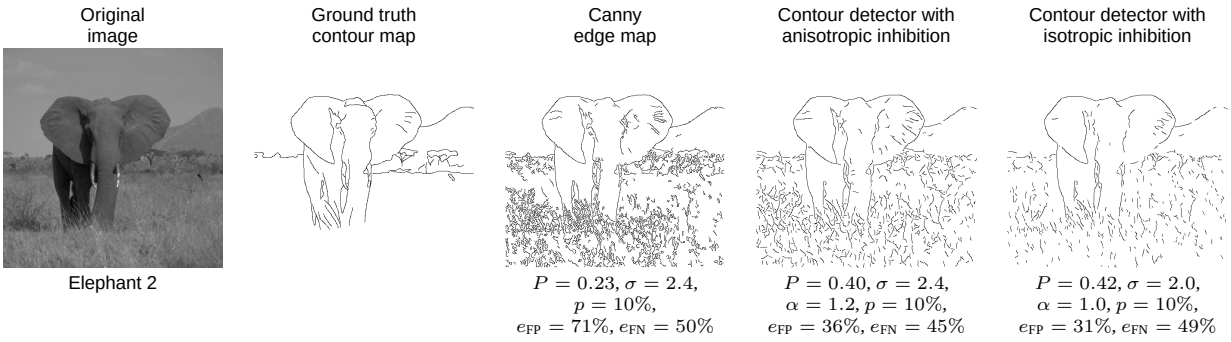


Fig. 4. From left to right: input image, corresponding ground truth contour map, the best contour map obtained with the Canny edge detector, the contour detector with anisotropic inhibition, and the contour detector with isotropic inhibition.

detector, we used eight scales with σ ranging from 1.0 to 2.4. For the contour detectors, we used four scales covering the same domain with σ ranging from 1.2 to 2.4, and two texture attenuation factors, $\alpha = 1.0$ and $\alpha = 1.2$. For all methods, we applied five high hysteresis threshold values, based on p ranging from $p = 50\%$ to $p = 10\%$. In total, we used 40 parameter combinations for each of the methods. We chose $N_\theta = 12$.

For similar values of σ and p, the contour maps delivered by the contour operators had better performance (up to a factor of two) than Canny's edge maps for all 40 images. This is mostly due to a reduced percentage of false positives, which is in agreement with the proposed model: edges resulting from a texture background (false positives) are suppressed, and object contours are retained. Fig. 4 shows the best performance edge maps for one of our test images.

4. DISCUSSION

Inhibition mechanisms have been applied previously to biologically motivated edge detectors in order to improve certain aspects of their function. A symmetric Gabor filter, will, for instance, respond not only along a line but also alongside the line at a certain distance from it. Similarly, the largest response of an antisymmetric Gabor filter to a line will be displaced from the line. In [7], various inhibition mechanisms have been proposed to remove these flanking responses. These works differ from the current work in two major aspects. First, the inhibition mechanisms act within the CRF. Second, the purpose of the inhibition is quite different: it deals with the removal of flanking responses, rather than with the suppression of texture edges.

The isotropic inhibition can be applied as an additional processing step to most edge detectors. More specifically, it can be added to the Canny edge detector as an intermediate step between the gradient computation and the edge thinning and binarization [6]. With this modification, the Canny algorithm becomes similar to the contour operator with isotropic inhibition presented above.

In summary, in this work we have shown that the (biologically inspired) non-CRF inhibition is a useful computational mechanism which reflects well human perception, and which can substantially improve the performance of contour detectors.

5. REFERENCES

[1] H. C. Nothdurft, J. L. Gallant, and D. C. van Essen, "Response modulation by texture surround in primate area V1: Correlates of "popout" under anasthesia," *Vis. Neuro.*, vol. 16, pp. 15–34, 1999.

[2] J. Canny, "A computational approach to edge detection," *IEEE Trans. PAMI*, vol. 8, no. 6, pp. 679–698, 1986.

[3] N. Petkov and M. A. Westenberg, "Suppression of contour perception by band-limited noise and its relation to non-classical receptive field inhibition," *Biol. Cyb.*, vol. 88, no. 3, pp. 236–246, 2002.

[4] J. G. Daugman, "Uncertainty relations for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters," *J. Opt. Soc. Am. A*, vol. 2, pp. 1160–1169, 1985.

[5] K. Bowyer, C. Kranenburg, and A. Dougherty, "Edge detector evaluation using empirical ROC curves," *CVIU*, vol. 84, pp. 77–103, 2001.

[6] C. Grigorescu, N. Petkov, and M. A. Westenberg, "Performance enhancement of contour detectors by surround inhibition," in *Int. Conf. Comp. Vis. Graph.*, Poland, September 25–29 2002, pp. 313–318.

[7] N. Petkov, P. Kruizinga, and T. Lourens, "Lateral inhibition in cortical filters," in *Proc. Int. Conf. on Dig. Signal Proc.*, Cyprus, July 14–16 1993, pp. 122–129.