View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by OAR@UM

Contour Detection by CORF Operator

George Azzopardi and Nicolai Petkov

Johann Bernoulli Institute for Mathematics and Computer Science,
University of Groningen, The Netherlands
{g.azzopardi,n.petkov}@rug.nl

Abstract. We propose a contour operator, called CORF, inspired by the properties of simple cells in visual cortex. It combines, by a weighted geometric mean, the blurred responses of difference-of-Gaussian operators that model cells in the lateral geniculate nucleus (LGN). An operator that has gained particular popularity as a computational model of a simple cell is based on a family of Gabor Functions (GFs). However, the GF operator short-cuts the LGN, and its effectiveness in contour detection tasks, which is assumed to be the primary biological role of simple cells, has never been compared with the effectiveness of alternative operators. We compare the performances of the CORF and the GF operators using the RuG and the Berkeley data sets of natural scenes with associated ground truths. The proposed CORF operator outperforms the GF operator (RuG: $t(39)=4.39$, $p<10^{-4}$ and Berkeley: $t(499)=4.95$, $p<10^{-6}$).

Keywords: Contour detection, Gabor function, LGN, simple cell, brain-inspired machine vision, orientation selectivity.

1 Introduction

Contour detection is believed to be the primary biological role of neurons in area V1 of visual cortex. There are two main classes of orientation-selective neurons in area V1, which have been named simple and complex cells by Hubel and Wiesel [10]. Typically, such a neuron would respond to an edge or a line of a given orientation in a given area of the visual field, called its receptive field (RF). The RFs of simple cells can be divided in excitatory and inhibitory regions.

The findings from the above research had instigated further electrophysiological studies and motivated various brain-inspired contour detectors. Two-dimensional (2D) Gabor functions (GFs) have been claimed to fit well the 2D RF profiles of cat simple cells [7,11] and have gained particular popularity in the field of image processing.

The computation of a Gabor filter is a linear spatial summation of the intensity values weighted by a Gabor function. The half-wave rectified response of a Gabor filter at a specific location in a given image models the activity of a simple cell. However, the GF model does not take into account the anatomical structure of the visual system as it uses as inputs the intensity values of an image as these are projected on the retina and bypasses the lateral geniculate nucleus (LGN) within the thalamus [5,8].

A.E.P. Villa et al. (Eds.): ICANN 2012, Part I, LNCS 7552, pp. 395–402, 2012.
© Springer-Verlag Berlin Heidelberg 2012

It has been speculated that the RF profile of a simple cell and its orientation selectivity are the result of a specific alignment of the RFs of LGN cells that provide input [10]. In this work we consider such a model and propose an operator, which we call Combination of Receptive Fields (CORF).

The local contour information that is extracted by orientation-selective cells provides basis for further, more complex visual tasks, such as object recognition [16,14,12,19,17,15]. The performance of various contour operators, which are inspired by the function of simple cells, in contour detection tasks has, however, not been quantified and they have not been compared in that respect. In the following, we evaluate and compare the performances of the proposed CORF operator and the GF operator, using two public data sets of images of natural scenes with associated contour ground truths.

The paper is organized as follows. In Section 2 we present the CORF operator. In Section 3 we evaluate the performance of the CORF operator in a contour detection task and compare it with that of the GF operator. We provide a discussion and draw conclusions in Section 4.

2 Combination Of Receptive Fields (CORF) Operator

Fig. 1 depicts an overview of the proposed CORF operator. Each of the light and dark disks in Fig. 1 represents the area of support (or RF) of a sub-unit that receives input from a pool of center-on ('+') or center-off ('-') center-surround operators, which model LGN cells. Such model LGN cells detect contrast changes. A sub-unit computes the sum of the weighted responses of a group of center-surround operators with the same polarity (on or off) and RF size. The RF of the sub-unit is the union of the RFs of the involved center-surround operators and it has the same polarity as these operators. In this way, a sub-unit detects contrast changes, similar to a model LGN cell, but it does so in a wider area.

A CORF operator achieves orientation selectivity by combining the responses of given sub-units with appropriate polarities and alignment of their RFs, Fig. 1. The operator parameters are determined in an automatic configuration process in which an example edge of a given orientation and polarity is presented. This pattern gives rise to a certain local configuration of center-surround operator responses in the area of support (or RF) of the concerned CORF operator. This local configuration is used to determine the polarity of the involved sub-units and their mutual spatial arrangement. The response of a CORF operator is computed

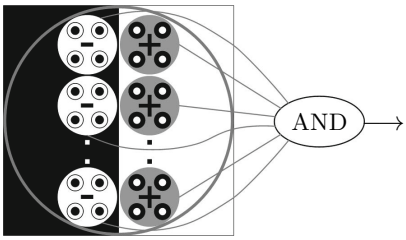


Fig. 1. Sketch of the proposed CORF operator. Orientation selectivity is achieved by combining the responses of two parallel sets of co-linear sub-units: one of center-on ('+') and the other of center-off ('-') type. Each sub-unit computes a sum of the weighted responses of a local group of center-surround operators.

as the weighted geometric mean of the sub-unit responses. In this way, a CORF operator generates a response only when all its afferent inputs are stimulated. In the following sub-sections we explain the configuration process in more detail.

2.1 Center-Surround Operators: Difference-of-Gaussians Functions

We use difference-of-Gaussians (DoG) functions with center-surround RFs, which we denote by DoG_σ^δ, the responses of which provide input to the sub-units mentioned above. The symbol δ represents the polarity (+ for center-on and - for center-off[1]) and the parameter σ is the standard deviation of the outer Gaussian function. We fix the standard deviation of the inner Gaussian function to 0.5σ. These functions are established models of LGN cells [6].

We denote by $c_\sigma^+(x,y)$ the response of a center-on DoG operator with a RF centered at image coordinates (x,y). It is computed by linear spatial summation of the intensity distribution $I(x',y')$ in the input image, weighted with the function $DoG_\sigma^+(x–x', y–y')$, followed by half-wave rectification. The response $c_\sigma^-(x,y)$ of a center-off DoG operator is the negative of $c_\sigma^+(x,y)$: $c_\sigma^-(x,y) = -c_\sigma^+(x,y)$.

2.2 Sub-unit and Its Parameters

Fig. 2 illustrates the process of configuring a CORF operator. The outer circle in Fig. 2b demarks the area of support (or RF) of that operator. Its center is positioned on an edge in the input image (Fig. 2a), which gives rise to DoG responses rendered in Fig. 2b. The eight small spots represent the RF centers of eight sub-units, four of center-on and four of center-off type. We determine such sub-units in a CORF operator as follows. We take two (in general k) concentric circles centered on the RF center of the concerned CORF operator and consider the DoG responses along these circles, Fig. 2c. The positions along these circles at which these responses reach significant local maxima are the positions at which we include sub-units. For the considered example, there are eight such positions, which result in the inclusion of eight sub-units in the operator at hand.

Each sub-unit is represented in parametric form by a tuple $(\delta, \sigma, \rho, \phi)$ where the parameters represent the polarity δ of the sub-unit, the scale parameter σ of the involved DoG functions in its pool, the radius ρ and the polar angle ϕ of the RF center of the sub-unit relative to the RF center of the CORF operator.

We denote by $S = \{(\delta_i, \sigma_i, \rho_i, \phi_i) \mid i = 1 \ldots n\}$ the set of 4-tuples that represent the configured sub-units above. For the concerned operator in Fig. 2, for $\sigma = 5$, and two configuration circles ($\rho \in \{18, 34\}$) for an image of size 100×100 pixels, the method described above results in eight sub-units, which are specified by the tuples in the following set: $S = \{\delta_1 = -, \sigma_1 = 5, \rho_1 = 34, \phi_1 = 1.48), (\delta_2 = +, \sigma_2 = 5, \rho_2 = 34, \phi_2 = 1.66), (\delta_3 = +, \sigma_3 = 5, \rho_3 = 34, \phi_3 = 4.62), (\delta_4 = -, \sigma_4 = 5, \rho_4 = 34, \phi_4 = 4.80), (\delta_5 = -, \sigma_5 = 5, \rho_5 = 18, \phi_5 = 1.41), (\delta_6 = +, \sigma_6 = 5, \rho_6 = 18, \phi_6 = 1.74), (\delta_7 = +, \sigma_7 = 5, \rho_7 = 18, \phi_7 = 4.55), (\delta_8 = -, \sigma_8 = 5, \rho_8 = 18, \phi_8 = 4.88)\}$.

[1] A center-on RF has its central region excitatory and its surrounding is inhibitory. A center-off RF has the opposite polarity.

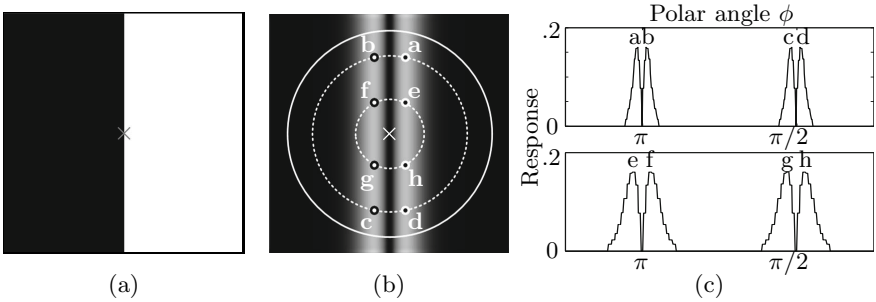


Fig. 2. Configuration of a CORF operator. (a) Synthetic edge of maximum contrast. (b) The gray-level rendering is the superposition of half-wave rectified responses of a center-on and a center-off DoG operators. The largest circle depicts the boundary of the area of support of the CORF operator, the center of which is illustrated by the '×' marker and lies on a vertical edge. (c) The top and bottom plots show the DoG responses along the outer and inner interrupted circles in (b), respectively. The labeled small spots represent the RF centers of eight sub-units, four center-on (white spots with black boundary) and four center-off (black spots with white boundary) with coordinates that are determined by the corresponding labeled local maxima in the plots in (c).

The first tuple in the set S, for instance, describes a sub-unit, which combines the responses of center-off ($\delta_1 = -$) DoG operators, with a RF size characterized by a standard deviation of ($\sigma_1 =$) 5 pixels, around a position at a radius of ($\rho_1 =$) 34 pixels and an angle of ($\phi_1 =$) 1.48 radians with respect to the RF center of the CORF operator; the RF center of this sub-unit is marked by 'a' in Fig. 2b.

2.3 Sub-unit Response

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

$$s_{\delta_i,\sigma_i,\rho_i,\phi_i}(x,y) \stackrel{\text{def}}{=} \sum_{x'} \sum_{y'} \{c^{\delta_i}_{\sigma_i}(x - \Delta x_i - x', y - \Delta y_i - y') G_{\sigma'}(x', y')\} \tag{1}$$

$$\Delta x_i = -\rho_i \cos\phi_i, \;\; \Delta y_i = -\rho_i \sin\phi_i, \;\; -3\sigma' \leq x', y' \leq 3\sigma'$$

The standard deviation σ' of the Gaussian function $G_{\sigma'}$ is a linear function[2] of the parameter ρ which is consistent with neurophysiological evidence for the relationship between the eccentricity and the average RF diameter of LGN cells [21]. Eq. 1 presents a convolution of the weighting function $G_{\sigma'}$ with the function $c^{\delta_i}_{\sigma_i}(x,y)$ that is shifted by the vector $(\Delta x_i, \Delta y_i)$, which is determined by the sub-unit parameters (ρ_i, ϕ_i).

[2] $\sigma' = (d_0 + \alpha\rho)/6$. We set $d_0 = 2$ and $\alpha = 0.9$.

2.4 CORF (Rotation-Invariant) Output

We define the response r_S of a CORF operator as the weighted geometric mean of the responses of all sub-units characterized by the tuples in the set S:

$$r_S(x,y) \stackrel{\text{def}}{=} \left(\prod_{i=1}^{|S|} \left(s_{\delta_i,\sigma_i,\rho_i,\phi_i}(x,y) \right)^{\omega_i} \right)^{1/\sum_{i=1}^{|S|}\omega_i} \tag{2}$$

where $\omega_i = \exp^{-\frac{\rho_i^2}{2\sigma'^2}}$ and $\sigma' = \frac{1}{3}\max_{i\in\{1\ldots|S|\}}\{\rho_i\}$. The multiplicative character of the above computation implies that the concerned CORF operator achieves a response only when all its afferent sub-units are active. It can be seen as a kind of soft AND function. The input contribution of sub-units decreases with an increasing distance of their RF centers from the RF center of the CORF operator. We are not aware of any anatomical or physiological evidence in support of this aspect of our model. It is rather our design decision to give a higher weight to afferent input that is closer to the RF center. Moreover, we obtained better results with this type of weighted multiplication than with a simple multiplication.

The orientation preference of a CORF operator as defined above depends on the orientation of the edge used for its configuration. One can create operators with different orientation preference by presenting different edges. Alternatively, one can manipulate the parameters in the set S, which corresponds to orientation preference for 0° to obtain a new set $\Re_\psi(S)$ with orientation preference ψ: $\Re_\psi(S) = \{(\delta_i,\sigma_i,\rho_i,\phi_i+\psi) \mid \forall\ (\delta_i,\sigma_i,\rho_i,\phi_i) \in S\}$. We then define a rotation invariant response as $\hat{r}_S(x,y) = \max_{\psi\in\Psi}\{r_{\Re_\psi(S)}(x,y)\}$, where Ψ is a set of n_θ equidistant orientations given as $\Psi = \{\frac{2\pi}{n_\theta}i \mid 0 \le i < n_\theta\}$.

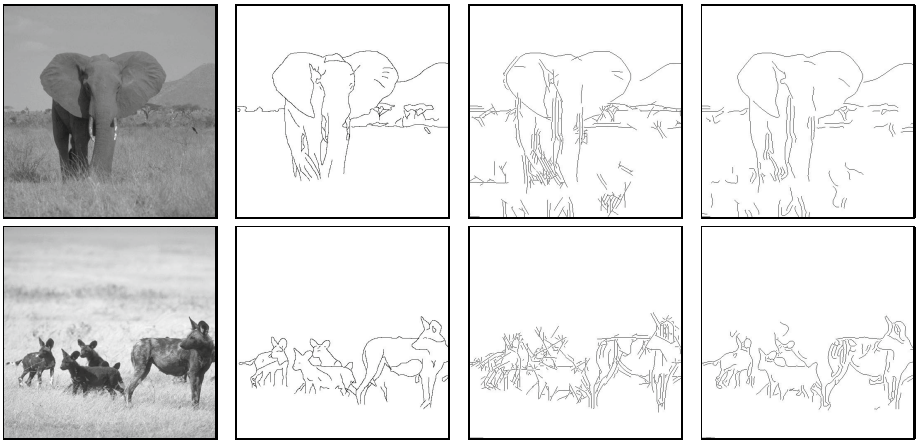
3 Results: CORF versus GF in Contour Detection

3.1 Evaluation Procedure

Fig. 3 (first column) shows two images of natural scenes and associated ground truth contour maps (second column) defined by a human observer. These images belong to the RuG data set[3], which was first used in [9] and comprises 40 images along with their corresponding ground truths. The ground truth images depict only the contours of objects, while the edges that are caused by texture are ommitted. We also use the Berkeley data set [1] of 500 images, which comprises a collection of (5 to 10) ground truth contour maps for every image.

To obtain a binary contour map for a given input image we apply to the output of the concerned operator (CORF or GF) a procedure widely used in computer vision. It consists of edge thinning by non-maxima suppression followed by binarization by hysteresis thresholding [4,20]. The former step essentially determines the ridges in the operator response. The latter step requires two parameter values, referred to as the low t_l and high t_h thresholds. We set $t_l =$

[3] The complete data set can be downloaded from `http://www.cs.rug.nl/~imaging`



(a) Input image (b) Ground truth (c) GF (d) CORF

Fig. 3. (a) Images of natural scenes from the RuG data set. (b) The corresponding contour maps designed by a human observer. Best contour maps obtained by (c) the GF and by (d) the proposed CORF operators, respectively.

$0.5t_h$. The resulting contour map depends on the t_h threshold value used: the lower that value, the larger the number of contour pixels in the map.

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

The values of R and P depend on the threshold value used for binarization: P increases and R decreases with an increasing value of t_h. We then compute the harmonic mean [13], $H = 2PR/(P+R)$, in order to come to a single performance measure that allows us to compare the results obtained by the two operators. We consider as an optimal result the operator output for which this harmonic mean reaches its maximum for a given image. We use the tolerance method proposed in [9] to evaluate inexact contour localization. The contour maps shown in Fig. 3c-d correspond to maximum values of the harmonic mean for the corresponding images. For the Berkeley data set we use the method suggested in [13] to compare an operator contour map with multiple ground truth contour maps.

3.2 Experimental Results

The CORF and the GF operators that we compare share two parameters: the number of orientations, n_θ, which we set to 12 (in intervals of $\pi/6$) and a scale parameter σ. The GF operator requires two other parameters, which we set as suggested in [18]: the wavelength $\lambda = \sigma/0.4$ and the spatial aspect ratio $\gamma = 0.5$. For every input image, we only consider the two maximum harmonic mean values, one for each of the two operators, across all nine scales ($\sigma \in \{1, 1.5, \ldots, 5\}$)

that we use. We configure a CORF operator with a number of ρ values[4] that depend on σ.

We apply a right-tailed paired-samples t-test to the set of pairs of harmonic mean values that are achieved by the two operators. The CORF operator outperforms the GF operator with high statistical confidence for both the RuG ($t(39)=4.39$, $p<10^{-4}$) and the Berkeley ($t(499)=4.95$, $p<10^{-6}$) data sets.

Notable is the fact that, compared to the GF operator, the best contour maps obtained by the proposed CORF operator, such as the ones shown in Fig. 3c-d contain less texture and the high curvature points are better preserved.

4 Discussion and Conclusions

We propose a novel operator for contour detection that is inspired by the anatomy of the visual system and the function of simple cells. It achieves orientation-selectivity by combining in an AND-type operation multiple DoG responses with properly aligned center-surround RFs. This is in contrast to the GF operator, which uses the intensity values of the given image as projected on the retina.

In [2] we show that the proposed CORF operator can also be considered as a computational model of a simple cell. There we show that the multiplicative character of the computation we use to combine model LGN responses is essential to achieve three important properties, namely cross-orientation suppression, contrast invariant orientation tuning and response saturation, which are typical of simple cells. In [3] we demonstrate that the CORF approach can also be used to model shape-selective neurons in area V4 of visual cortex.

The implementation of the proposed CORF operator is rather straightforward: it includes blurring (achieved by convolving with a Gaussian function) of half-wave rectified responses of DoG operators, shifting appropriately these blurred responses by different vectors, which are determined in the configuration of the operator, and using them for the pixel-wise evaluation of a weighted geometric mean that gives the output of the CORF operator.

We demonstrated that the proposed CORF operator is more effective than the GF operator in contour detection, which is assumed to be the primary biological role of simple cells. The improvement is mainly attributable to robustness to noise, better edge localization and better preservation of curvatures.

References

1. Arbelaez, P., Maire, M., Fowlkes, C., Malik, J.: Contour detection and hierarchical image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 33(5), 898–916 (2011)
2. Azzopardi, G., Petkov, N.: A CORF computational model of a simple cell that relies on LGN input outperforms the Gabor function model. Biological Cybernetics 106(3), 177–189 (2012)

[4] For $\sigma \in \{1, 1.5, 2\}$ we use three radii ($\rho \in \{3, 7, 14\}$), for $\sigma \in \{2.5, 3, 3.5\}$ we use four radii ($\rho \in \{3, 6, 13, 25\}$) and for $\sigma \in \{4, 4.5, 5\}$ we use 5 radii ($\rho \in \{3, 5, 9, 18, 34\}$).

3. Azzopardi, G., Petkov, N.: Trainable COSFIRE Filters for Keypoint Detection and Pattern Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence (preprint, 2012)
4. Canny, J.: A computational approach to edge-detection. IEEE Transactions on Pattern Analysis and Machine Intelligence 8(6), 679–698 (1986)
5. Chung, S., Ferster, D.: Strength and orientation tuning of the thalamic input to simple cells revealed by electrically evoked cortical suppression. Neuron 20(6), 1177–1189 (1998)
6. Croner, L.J., Kaplan, E.: Receptive-fields of P-ganglion and M-ganglion cells across the primate retina. Vision Research 35(1), 7–24 (1995)
7. Daugman, J.G.: Uncertainty relation for resolution in space, spatial-frequency, and orientation optimized by two-dimensional visual cortical filters. Journal of the Optical Society of America a-Optics Image Science and Vision 2(7), 1160–1169 (1985)
8. Ferster, D., Chung, S., Wheat, H.: Orientation selectivity of thalamic input to simple cells of cat visual cortex. Nature 380(6571), 249–252 (1996)
9. Grigorescu, C., Petkov, N., Westenberg, M.A.: Contour detection based on nonclassical receptive field inhibition. IEEE Transactions on Image Processing 12(7), 729–739 (2003)
10. Hubel, D.H., Wiesel, T.N.: Receptive fields, binocular interaction and functional architecture in cats visual cortex. J. Physiol. (Lond.) 160(1), 106–154 (1962)
11. Jones, J.P., Palmer, L.A.: An evaluation of the two-dimensional gabor filter model of simple receptive-fields in cat striate cortex. Journal of Neurophysiology 58(6), 1233–1258 (1987)
12. Kovesi, P.: Image features from phase congruency. Videre—Videre 1(3) (1999)
13. Martin, D.R., Fowlkes, C.C., Malik, J.: Learning to detect natural image boundaries using local brightness, color, and texture cues. IEEE Transactions on Pattern Analysis and Machine Intelligence 26(5), 530–549 (2004)
14. Mehrotra, R., Namuduri, K.R., Ranganathan, N.: Gabor filter-based edge-detection. Pattern Recognition 25(12), 1479–1494 (1992)
15. Meyers, E., Wolf, L.: Using biologically inspired features for face processing. International Journal of Computer Vision 76(1), 93–104 (2008)
16. Morrone, M.C., Burr, D.C.: Feature detection in human-vision - a phase-dependent energy-model. Proceedings of the Royal Society of London Series B-Biological Sciences 235(1280), 221–245 (1988)
17. Mutch, J., Lowe, D.G.: Object class recognition and localization using sparse features with limited receptive fields. International Journal of Computer Vision 80(1), 45–57 (2008); 24th Conference on Computer Vision and Pattern Recognition, New York, NY, June 17-22 (2006)
18. Petkov, N.: Biologically motivated computationally intensive approaches to image pattern-recognition. Future Generation Computer Systems 11(4-5), 451–465 (1995)
19. Serre, T., Wolf, L., Bileschi, S., Riesenhuber, M., Poggio, T.: Robust object recognition with cortex-like mechanisms. IEEE Transactions on Pattern Analysis and Machine Intelligence 29(3), 411–426 (2007)
20. Sonka, M., Hlavac, V., Boyle, R.: Image processing, analysis, and machine vision. Brooks/Cole, Pacific Grove (1999)
21. Xu, X.M., Ichida, J.M., Allison, J.D., Boyd, J.D., Bonds, A.B., Casagrande, V.: A comparison of koniocellular, magnocellular and parvocellular receptive field properties in the lateral geniculate nucleus of the owl monkey (Aotus trivirgatus). Journal of Physiology-London 531(1), 203–218 (2001)

View publication stats