

University of Groningen

A robust contour detection operator with combined push-pull inhibition and surround suppression

Melotti, Damiano; Heimbach, Kevin; Rodriguez-Sanchez, Antonio; Strisciuglio, Nicola; Azzopardi, George

Published in:
Information Sciences

DOI:
[10.1016/j.ins.2020.03.026](https://doi.org/10.1016/j.ins.2020.03.026)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2020

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Melotti, D., Heimbach, K., Rodriguez-Sanchez, A., Strisciuglio, N., & Azzopardi, G. (2020). A robust contour detection operator with combined push-pull inhibition and surround suppression. *Information Sciences*, 524, 229-240. <https://doi.org/10.1016/j.ins.2020.03.026>

Copyright

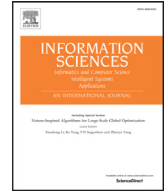
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.



A robust contour detection operator with combined push-pull inhibition and surround suppression



Damiano Melotti^{a,b}, Kevin Heimbach^c, Antonio Rodríguez-Sánchez^d, Nicola Strisciuglio^e, George Azzopardi^{a,*}

^a Nijenborgh 9, University of Groningen, Groningen, 9747AG, the Netherlands

^b Via Sommarive 9, DISI, University of Trento, 38123, Italy

^c Msida MSD2080, University of Malta, Malta

^d Technikerstrasse 21a, University of Innsbruck, A-6020, Austria

^e Hallenweg 19, University of Twente, Enschede, 7522NH, the Netherlands

ARTICLE INFO

Article history:

Received 17 July 2019

Revised 7 March 2020

Accepted 11 March 2020

Available online 17 March 2020

Keywords:

Contour detection

Simple cell

Push-pull inhibition

Surround suppression

ABSTRACT

Contour detection is a salient operation in many computer vision applications as it extracts features that are important for distinguishing objects in scenes. It is believed to be a primary role of simple cells in visual cortex of the mammalian brain. Many of such cells receive push-pull inhibition or surround suppression. We propose a computational model that exhibits a combination of these two phenomena. It is based on two existing models, which have been proven to be very effective for contour detection. In particular, we introduce a brain-inspired contour operator that combines push-pull and surround inhibition. It turns out that this combination results in a more effective contour detector, which suppresses texture while keeping the strongest responses to lines and edges, when compared to existing models. The proposed model consists of a Combination of Receptive Field (or CORF) model with push-pull inhibition, extended with surround suppression. We demonstrate the effectiveness of the proposed approach on the RuG and Berkeley benchmark data sets of 40 and 500 images, respectively. The proposed push-pull CORF operator with surround suppression outperforms the one without suppression with high statistical significance.

© 2020 The Author(s). Published by Elsevier Inc.
This is an open access article under the CC BY-NC-ND license.
(<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

1. Introduction

Contour detection is a very important tool in computer vision as it allows the reduction of images to salient features. Many contour detection algorithms suffer from insufficient ability to distinguish edges that are embedded in texture, such as grass and leaves, from those that delineate the defining features of objects. This distinction of salient edges is, however, a seamless operation to the mammalian brain.

The main source of inspiration for this work is neuro-physiological evidence in order to design a contour detection operator that exhibits high robustness to texture and accentuates contours that characterize objects in complex scenes.

* Corresponding author.

E-mail address: g.azzopardi@rug.nl (G. Azzopardi).

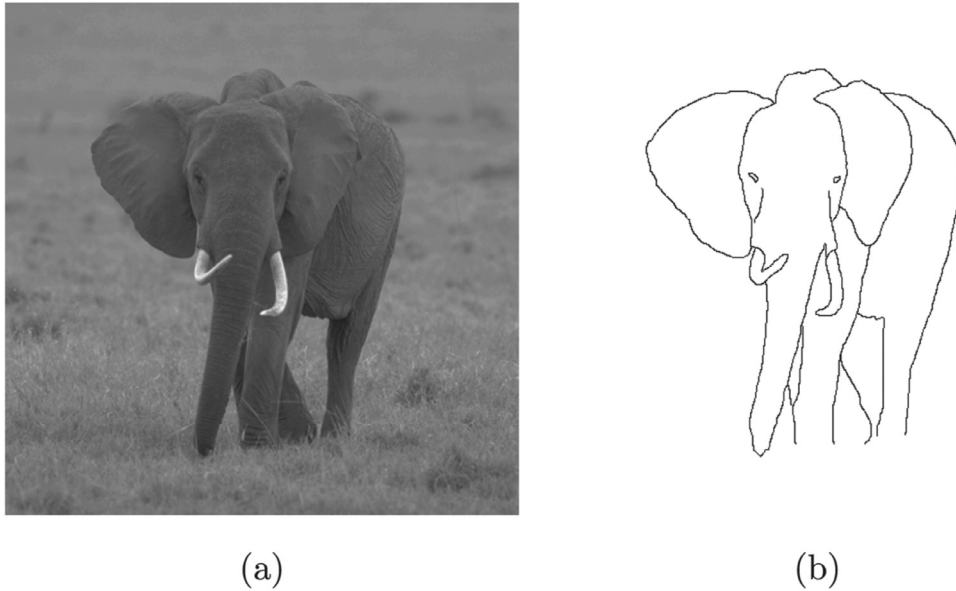


Fig. 1. (a) An example image of a real-life scene (taken from the RuG data set) with a main object surrounded by texture. (b) The desired output contour map hand drawn by a human being.

Fig. 1a shows an example of an image with a dominant salient object surrounded by texture in the form of grass. **Fig. 1b** shows the desired contour map that delineates only the boundary and the salient features within the main object.

The aim of this work is to design a brain-inspired local operator that emphasizes the salient contours that humans perceive as whole objects and suppresses non-relevant responses. Next, we provide a brief summary on how the brain processes visual information and describe two inhibition phenomena manifested by certain types of cells, which we mimic in order to design our operator.

The ventral and dorsal streams in the visual system of the mammalian brain are the areas where object recognition and motion detection are, respectively, believed to be performed. The primary visual cortex (V1) is the first area of visual information processing in these streams. It contains, among other types of neurons, simple cells that respond to lines or edges with preferred orientations [1]. The primary biological function of simple cells is believed to be edge detection, which is the first step towards contour detection and object recognition. In [2], Hubel and Wiesel speculated that the preferred stimulus of a simple cell depends on the specific alignment of the concentric receptive fields of multiple LGN cells that provide input to it. A simple cell responds when all afferent LGN cells are activated, meaning that the preferred stimulus lies within its receptive field.

The receptive field of a simple cell is organized into inhibitory and excitatory regions. This is known as the classical receptive field (CRF). There are, however, simple cells that receive inhibition from other such cells whose receptive fields lie in the surroundings of their receptive fields. These types of cells are known to have non-classical receptive fields (NCRF) [3]. Moreover, many simple cells receive antiphase or push-pull inhibition [4–7]. A push-pull response of a simple cell with a CRF is observed when two stimuli of preferred orientation but of opposite contrast evoke responses of opposite sign; the stimulus of preferred contrast evokes a push (positive) response and the stimulus of opposite contrast evokes a pull (negative) response (see Fig. 2). Although there is no explicit biological evidence for the wiring depicted in Fig. 2, the model is generally accepted in neurophysiology [4,5,8–11]. It is likely that the push-pull inhibition mechanism contributes to the suppression of texture, thus giving more prominence to the contour of objects. It seems that it is also the most dominant form of inhibition received by simple cells [5–7,12].

The field of computer vision has often been inspired by neurophysiological findings. Simple cells and their mechanisms have also been modeled by computer algorithms and used for contour detection [14–16]. Originally, simple cells were modeled by 2D Gabor functions and have been widely used in computer vision applications [17,18]. It turns out, however, that Gabor functions do not explain all properties of simple cells and they also bypass the intermediate responses of model LGN cells. This aspect was the focus of a recent study [19], which proposed a novel Combination of Receptive Fields (CORF) model of a simple cell. It was demonstrated that the new CORF model achieves more properties of real simple cells, such as contrast invariant orientation tuning, cross orientation suppression, and better contour detection, than the Gabor function model. Its implementation, called (B-)COSFIRE (Combination of Shifted Filter Responses), has been demonstrated to be successful for the delineation of elongated structures, such as blood vessels in retinal images [20], roads and rivers in aerial images [21] and pavement cracks [22]. Later, the CORF model was extended by incorporating an inhibitory step, in order to model push-pull inhibition [13]. It achieved further properties of simple cells and the performance of contour detection

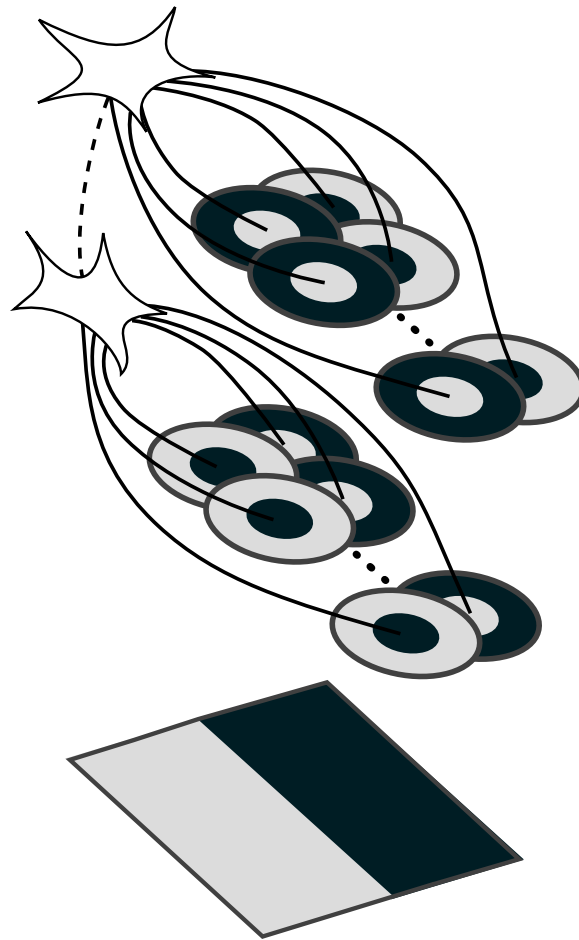


Fig. 2. A schematic diagram of a push-pull CORF model. The concentric circles represent the receptive fields of model LGN cells. The shaded light and dark gray areas indicate ON and OFF subregions. The set of concentric circles at the top can be considered as the model LGN cells that simultaneously fire for the preferred stimulus shown at the bottom, thus activating the simple cell they are connected to. The other concentric circles in the middle have opposite polarity and are connected to an intracortical model cell. The final response of the push-pull CORF model is the difference between the response of the model cell at the top and a weighted response of the model cell in the middle. Adapted from Azzopardi et al. [13].

improved substantially, mainly due to increased robustness to noisy contours. It also outperformed other contour detection models inspired by the brain as well as other classical detectors, such as the Canny operator. The effects of push-pull inhibition to strengthen the robustness to various types of noise and textured background have been recently shown in [23], where a novel robust inhibition-augmented curvilinear operator, named RUSTICO, has been proposed.

In the visual system of the brain there are also simple cells that exhibit NCRF inhibition, whose responses rely on the preferred stimuli in their excitatory regions but also on stimuli that lie in the surroundings of their receptive fields [24]. In the current context it means that the response to an oriented stimulus, such as a line, can be influenced by the presence of similar stimuli in its neighbourhood. This can manifest itself by a decreased response to a contour in the presence of surrounding texture [25]. The source of this mechanism and specifically whether it has an intracortical origin, however, remains a matter of debate [26].

The phenomenon of surround suppression has already been modeled and used for contour detection. In [27,28], the authors added a computational step of surround suppression to the Canny edge detector and to a Gabor-based contour operator. Their resulting operators responded strongly to isolated lines and edges, region boundaries and object contours, and exhibited weaker or no responses to texture. Fig. 3 illustrates a sketch of the non-classical receptive field of that model. The response of an edge detector that operates inside the white surface area is suppressed by the responses of the edge detectors operating in the surrounding gray surface area. In [29], another contour detection model with surround suppression was proposed. It contains a butterfly-shaped surrounding area for the inhibition component, and only one side sub-region that produces less inhibition contributes to the response of the operator.

The contribution of this work is three-fold. First, we propose a novel operator for contour detection that combines push-pull inhibition and surround suppression. The proposed model takes as input the response of a push-pull CORF model of a

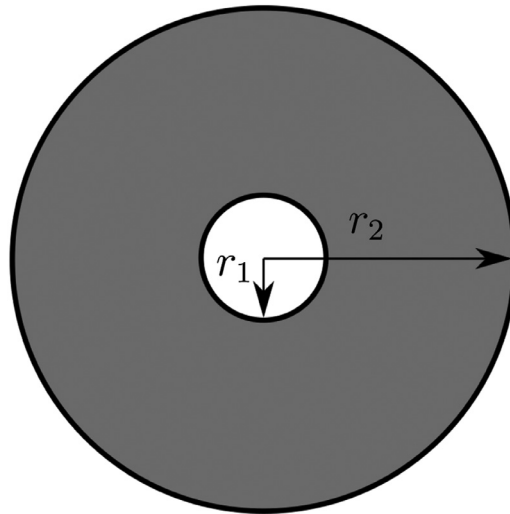


Fig. 3. The inner surface area with radius r_1 can be considered as the receptive field or support of an edge detection operator. The surround suppression originates from the outer gray surface area with inner radius r_1 and outer radius r_2 . Adapted from Grigorescu et al. [28].

simple cell [13], extended with a surround suppression inhibition step [28]. Second, we explore two strategies of surround suppression. Finally, we evaluate the performance of the proposed operator on the RuG and Berkeley benchmark data sets, composed of 40 and 500 images of natural scenes, respectively. We demonstrate with statistical analysis that the proposed operator with push pull inhibition and surround suppression improves the performance significantly.

The paper is organized as follows. In Section 2, we present the CORF model with push-pull inhibition and surround suppression. In Section 3, we report the experimental results obtained on the RuG and Berkeley data sets. In Section 4, we discuss the important elements of the model and its performance. Finally, we draw conclusions in Section 5.

2. Methods

2.1. Overview

In Fig. 4 we illustrate the idea that served as basis for designing the proposed operator: we combine the CORF push-pull inhibition in the inner part with a surround suppression mechanism in the outer part. The model can be considered as the receptive field of a simple cell in V1: the white surface area of the inner circle acts as the CRF region of this model cell. It takes as input the response of a CORF model with push-pull inhibition [19]. The outer circle with gray surface corresponds to the area where isotropic surround suppression is computed. This type of surround suppression means that the edges in the surrounding area contribute to response suppression of the concerned model cell, irrespective of their orientation.

2.2. Push-pull CORF model

We denote by $C_{\sigma,k,b}(x, y)$ the response of a push-pull CORF model at position (x, y) . It takes as input the responses of two CORF models (without inhibition) of the type proposed in [19]. Each afferent CORF model has appropriately aligned center-surround receptive fields and combines their output with a weighted geometric mean. The parameter σ is the outer standard deviation of the involved DoG functions with center-surround receptive fields. The two afferent CORF models are of opposite polarity, resulting in an excitatory (push) component and an inhibitory (pull) one. The responses of these components are combined by first multiplying the pull response with the given k parameter value and then subtract the result from the push response. The parameter b ($b > 0$) is used for the inhibitory component in order to produce a receptive field that is broader than that of the excitatory component. The orientation bandwidth of the pull model increases with an increasing value of b up to some extent. This property is supported by neurophysiological evidence [30,31].

2.3. Surround suppression

We extend the push-pull CORF model with a term that considers the surrounding area as another source of inhibition to the point of interest. We denote by $DoG_\gamma(x, y)$ a difference-of-Gaussians function:

$$DoG_\gamma(x, y) = \frac{1}{2\pi(4\gamma)^2} \exp\left(-\frac{x^2 + y^2}{2(4\gamma)^2}\right) - \frac{1}{2\pi(\gamma)^2} \exp\left(-\frac{x^2 + y^2}{2(\gamma)^2}\right), \quad (1)$$

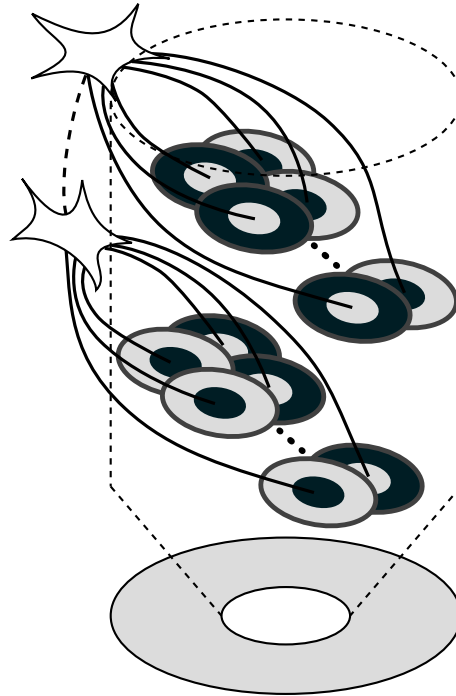


Fig. 4. The proposed CORF model with push-pull inhibition and surround suppression. The push-pull CORF operator (top) acts as the CRF of the current model. The surround suppression (bottom) area takes as input the combined responses of many push-pull CORF models in the shaded region.

where the two Gaussian functions are centered (i.e. their means are 0), and where γ is the size of the inner Gaussian function. Moreover, we use the same weighting function $w_\gamma(x, y)$ that was proposed in [27]:

$$w_\gamma(x, y) = \frac{\Theta(\text{DoG}_\gamma(x, y))}{\|\Theta(\text{DoG}_\gamma(x, y))\|_1}, \tag{2}$$

where Θ is the Heaviside function, which we use for rectification, and $\|\cdot\|_1$ is the L_1 norm.

For a given location (x, y) in an image a surround suppression term is computed. This term is a weighted sum of the responses of push-pull CORF models in the suppression surround of the concerned point (Fig. 4).

2.4. Isotropic surround suppression

We implement a form of surround suppression that does not take into account the orientation of surrounding edges, called isotropic surround suppression. In this type of suppression, only the distance to these edges is taken into account. The suppression term $s_{\gamma,\sigma,k,b}(x, y)$ is defined as the convolution of the response map of the push-pull CORF model $C_{\sigma,k,b}(x, y)$ with the weighting function $w_\gamma(x, y)$:

$$s_{\gamma,\sigma,k,b}(x, y) = \sum_{x'} \sum_{y'} w_\gamma(x - x', y - y') C_{\sigma,k,b}(x, y). \tag{3}$$

We define a contour operator $J_{\gamma,\sigma,k_1,k_2,b}$ that takes as input the responses of the push-pull CORF model $C_{\sigma,k,b}(x, y)$ and the isotropic suppression term $s_{\gamma,\sigma,k,b}(x, y)$ as:

$$J_{\gamma,\sigma,k_1,k_2,b}(x, y) = \Theta(C_{\sigma,k=k_1,b}(x, y) - \alpha s_{\gamma,\sigma,k=k_2,b}(x, y)). \tag{4}$$

The parameter α ($\alpha > 0$) indicates the suppression strength that the surround exerts on the push-pull CORF response. The way we define the operator J allows the flexibility to be used in two ways. When $k_1 = k_2$ and they are both greater than zero, the same type of CORF models with push-pull inhibition are deployed within and in the surrounding area of the receptive field. If $k_1 > 0$ and $k_2 = 0$, the responses of a CORF model with push-pull inhibition in the inner area are suppressed by the responses of CORF models *without* push-pull inhibition operating in the surrounding.

2.5. Thinning and hysteresis thresholding

For the binarization of the response of the proposed operator we use the two-step procedure that was proposed in [32]. First, the edges in the output image are thinned by non-maximum suppression to obtain the ridges. Subsequently,

hysteresis thresholding is applied to obtain a binary contour map. It requires a high threshold ζ , and a low threshold which we set to 0.5ζ , as suggested in [19]. Hysteresis thresholding is a widely used operation in the image processing literature concerning contour detection and other applications, and it works as follows. The pixels with values higher than ζ are retained and the pixels with values smaller than the low threshold are set to zero. The pixels whose values are within the low and high thresholds are kept only if they are connected to pixels whose responses are higher than ζ . The connection is determined through a link of other pixels with values larger than the low threshold. Similar to other works, in our experiments we vary the value of the high threshold ζ systematically from 0.1 to 1 in intervals of 0.1. The selectivity of the operator increases with increasing value of ζ , as it retains only the pixels with the strongest responses.

3. Experimental evaluation

3.1. Data sets

We tested the proposed contour detection operator on the RuG and Berkeley benchmark data sets. The RuG data set [27] was introduced for the evaluation of the Gabor function model with NCRF. It consists of 40 natural images of size 512×512 pixels. Each image is coupled with a hand drawn ground truth image, which contains only the most important contours of objects excluding texture and contours that are less sharp.

The Berkeley data set [33] is composed of 500 images (of size 481×321 or 321×481 pixels) of objects in complex scenes, that have been manually segmented by five different persons. This data set was mainly developed for the evaluation of image segmentation, but it has been widely used for developing and benchmarking contour detection algorithms. Fig. 5a contains examples from these two data sets, while Fig. 5b contains the corresponding hand drawn ground truth binary images.

3.2. Quantitative performance measure

In order to evaluate the performance of the proposed approach, we compute the Matthews Correlation Coefficient (MCC), which is a balanced measure of the classification accuracy also in cases where the positive and negative classes have a very unbalanced cardinality. The MCC has been previously used for performance evaluation of contour detection [13]. It is computed from the values of the confusion matrix as:

$$\text{MCC} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}, \quad (5)$$

where TP and FP stand for the number of true and false positives, and FN and TN stand for the number of true and false negatives. A TP occurs when the operator detects an edge pixel where the human-drawn ground truth also contains an edge. An FP occurs when an edge is detected where the ground truth does not contain an edge. A TN means that the operator does not detect an edge where it should not, i.e. the ground truth also did not contain an edge. An FN occurs when the operator fails to detect an edge where the ground truth marks an edge pixel.

As the ground truth might not be precise due to inaccuracy introduced by the person providing it, we use an evaluation method that takes into account near correct edge detection when calculating the MCC [28,34]. Using such tolerance, a detected edge is considered correct if the corresponding ground truth edge is within a 5×5 pixel area around the detected edge; i.e. at most 2 pixels away. Any given ground truth edge pixel can only be matched with one detected edge pixel, which leads to every edge pixel in the ground truth to be considered only once.

Since the Berkeley data set includes more than one ground truth per image, we calculate the MCC as a function of the total number of TPs, FPs, FNs, and TNs with respect to all five ground truth maps. This is based on what was proposed in [34] for computing the harmonic mean of precision and recall.

An MCC of 1 occurs when the output of the operator perfectly matches the hand drawn ground truth. An MCC of -1 indicates that the contour map of the operator is the inverted image of the hand drawn ground truth. And an MCC of 0 indicates random detection of edges.

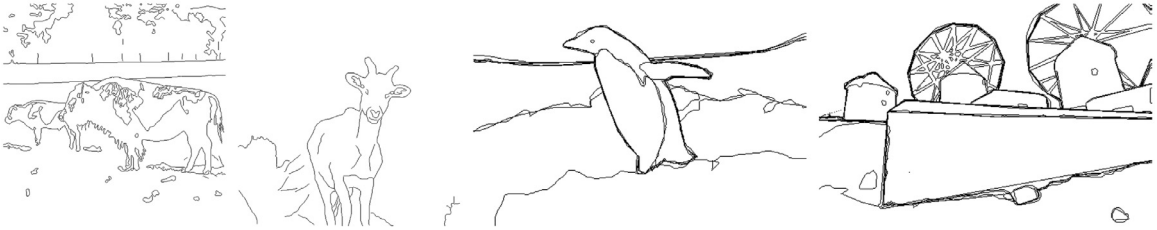
3.3. Evaluating model parameters

The model that we propose has five parameters, namely the standard deviation σ of the difference-of-Gaussians function involved in the push-pull CORF model component, the receptive field size factor b of the inhibitory component of the push-pull CORF model with respect to the excitatory component, the pull strength k of the inhibitory component of the push-pull CORF model, the inhibition strength α of the surround suppression component, and the inner circle size given by the standard deviation γ of the surround suppression component. The parameters k_1 and k_2 indicate whether the surround suppression is calculated based on a CORF model with or without push-pull inhibition. The outer circle size of the surround suppression component is fixed at 4 times the inner circle size γ as suggested in [28].

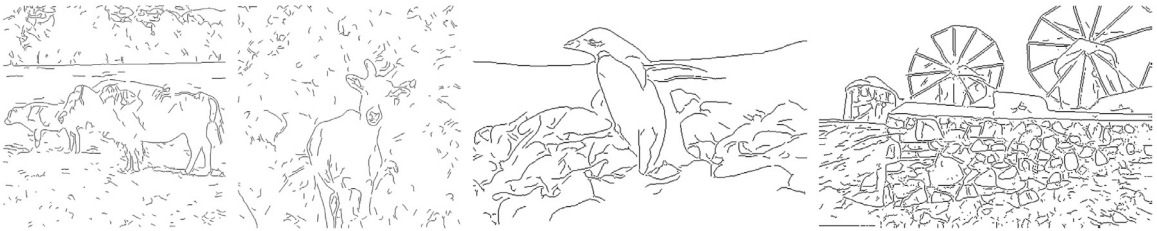
For the push-pull CORF model we use the same parameter values ($\sigma = 2.2$, $\beta = 4$, $k = 1.8$) which provided the best results in [13] on the RuG data set. We experiment with different combinations of values of the new hyperparameters



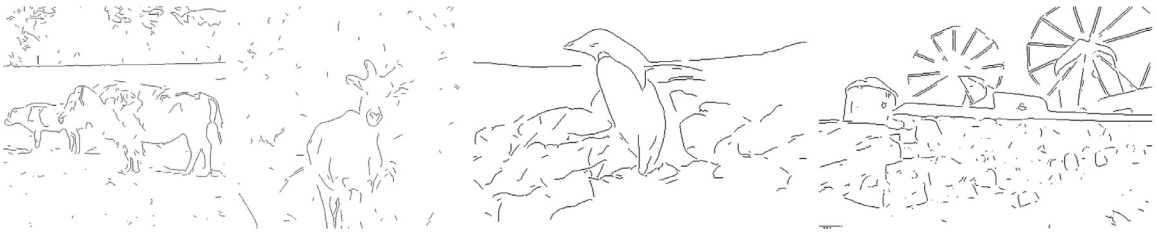
(a) Original images



(b) Ground truth contour maps



(c) PP only



(d) PP+SS: $k_1 = 1.8, k_2 = 0$

Fig. 5. (a) Examples of images taken from the RuG (first two images) and the Berkeley (last two images) data sets. (b) The corresponding hand drawn contour maps. For the Berkeley images the ground truth is rendered as the superimposition of binary ground truth contour maps hand drawn by five different persons. (c) The contour maps achieved by the CORF model with only push-pull inhibition, and (d) the output by the new push-pull CORF model with surround suppression.

(α and γ) of the surround suppression component. In order to avoid deviating from the scope of this work, we do not strive to determine the best set of all hyperparameters, as our goal is to test the hypothesis that the combination of both types of inhibition contributes to better contour detection performance than a CORF model with only push-pull inhibition.

The values considered for the parameters of the surround suppression component are $\gamma = \{2, 3\}$ and $\alpha = \{0.5, 1\}$. These variations are explored in the case when $k_1 = k_2 = 1.8$ (push-pull in the inner and outer circles of the surround suppression component), and in the case when $k_1 = 1.8$ and $k_2 = 0$ (push-pull only in the inner circle of the surround suppression component).

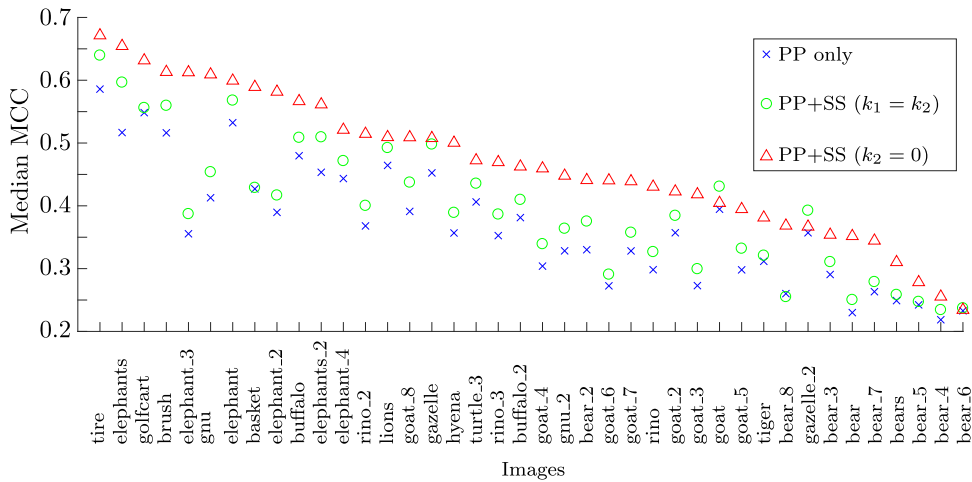


Fig. 6. Comparison of the median MCC values achieved by the three considered approaches on the RuG data set. The abbreviation PP stands for push-pull and SS stands for surround suppression. PP+SS refers to the proposed model that uses both push-pull inhibition and surround suppression. The labels in the x -axis are the names of the images in the RuG data set.

3.4. Results

For each image in the considered data sets, we first compute the thinned response map of a given contour operator and then we apply hysteresis thresholding by varying the high threshold ζ between 0.1 and 1 in steps of 0.1. The numbers of FPs, FNs and TNs increase and the number of TPs decreases with an increasing value of the parameter ζ . The MCC, therefore, improves as the ζ value increases up to the extent where the number of incorrect matches exceeds the number of correct ones.

We compute the MCC for each threshold value and combination of parameter values. This results in 40 ($[2 \text{ values of } \gamma] \times [2 \text{ values of } \alpha] \times [10 \text{ values of } \zeta]$) MCC values computed for each image when using the proposed operator with surround suppression configured with different sets of parameters. For the model without surround suppression we compute one MCC value for each threshold, resulting in 10 MCC values per image. In Fig. 6 we show a comparison of the median MCC values achieved for each image of the RuG data set by the proposed operator with surround suppression, configured to use push-pull inhibition in the center only ($k_2 = 0$) and also in the surround ($k_1 = k_2$), and by the CORF operator with push-pull inhibition only (i.e. no surround suppression). Moreover, in Fig. 7 we use box plots to compare the results achieved by the concerned operators on the same data set.

We achieved similar results on the Berkeley data set of 500 images. The proposed operator with push-pull and surround suppression outperforms the one without surround suppression in the majority of the cases. In Fig. 8, we show the box plots of the MCC values of 40 images, for both configurations of the model. The maximum 40 median MCC values achieved by the configuration ($k_1 = 1.8$, $k_2 = 0$) were used to select the 40 images indicated in Fig. 8.

Both contour detection models with surround suppression that we propose perform better than the push-pull CORF model without surround suppression. The last two rows of Fig. 5 show the resulting contour maps of four test images obtained by the proposed operator. It is evident that the contour maps obtained by the new model have less texture and more detail of the main objects of interest. We applied a paired-sample t -test to the set of pairs of MCC values that were achieved by the two models. The new model outperforms the previous model with high statistical significance both on the RuG data set ($k_1 = k_2$: $t(39) = -10.53$, $p < 0.001$; $k_2 = 0$: $t(39) = -10.92$, $p < 0.001$) and on the Berkeley data set ($k_1 = k_2$: $t(499) = -22.86$, $p < 0.001$; $k_2 = 0$: $t(499) = -14.83$, $p < 0.001$).

4. Discussion

In this paper, we show how the performance to contour detection of the existing push-pull CORF model [13] is substantially improved by incorporating surround suppression. In our experiments, we demonstrate that this additional step improves the resulting contour maps, with high statistical significance. The improvement is especially visible in images containing large amounts of texture. For instance, images of animals surrounded by natural landscapes that include grass and trees experience the largest improvement. In images where there is no or little texture, the improvement is negligible. In [13], it was shown that a CORF model with push-pull inhibition and without surround suppression outperformed several established contour detection operators, including Canny, and the Gabor function model with isotropic and anisotropic inhibition.

We carried out experiments with two strategies of surround suppression: one where the surround is characterized by CORF models without push-pull inhibition ($k_2 = 0$) and the other where the surround is characterized by CORF models

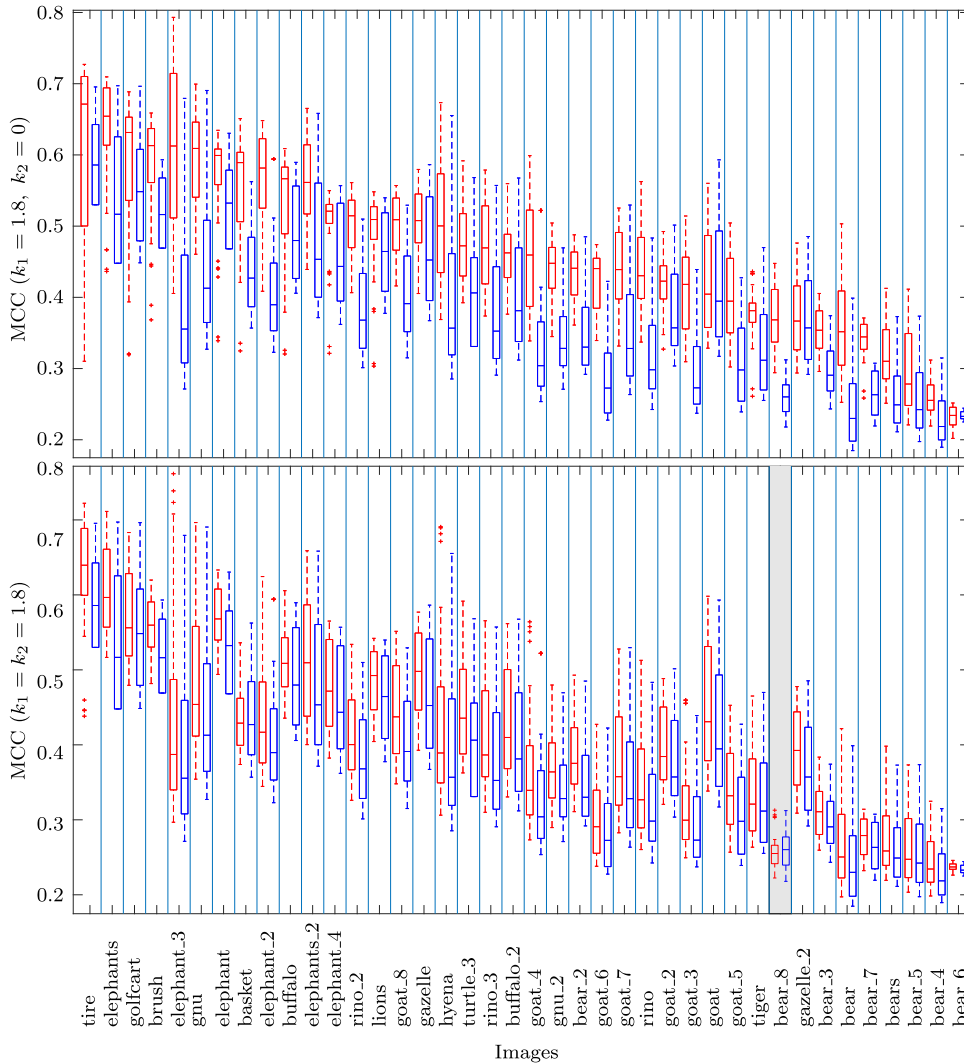


Fig. 7. Box plots of the performance of the new operator against the push-pull CORF model without surround suppression on the RuG data set. The distribution of scores for all possible parameter combinations of an operator are represented in one box plot per image. The values of the PP only model (blue) are the same for the top and bottom plots. The red box plots are generated from the results obtained by the proposed operator used in two ways: (top) $k_1 = k_2 = 1.8$ and (bottom) $k_1 = 1.8$ and $k_2 = 0$. The shaded column indicates the instance where the PP only model performs better than the proposed model. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

with push-pull inhibition ($k_1 = k_2 = 1.8$). For the RuG data set of 40 images both strategies achieved comparable statistical significance when compared to the operator without surround suppression, and for the Berkeley data set of 500 images, the latter approach achieved the best performance.

In images where there is a substantial amount of texture in the background - like the images in the RuG data set - are better processed by the proposed model when $k_2 = 0$. This happens because a denser response map is subtracted from the excitatory one, when no push-pull inhibition is used in the surround. The result is higher suppression of texture and therefore better quality of the final contour map. The reason why the three images, namely goat, gazelle_2, and bear_6 of the RuG data set, are better processed when $k_1 = k_2$ is due to the scale of the afferent difference-of-Gaussians functions. In our experiments, we used the set of hyperparameters, including $\sigma = 2.2$, that contributed to the best performance in [13]. When evaluating with a set of scale values ($\sigma \in \{1.5, 2.2, 2.9, 3.5\}$) the median MCC values obtained with $k_2 = 0$ are all better than those obtained with $k_1 = k_2 = 1.8$. We do not elaborate further on this aspect as the scope of this work is to test the hypothesis that adding surround suppression to push-pull inhibition results in a contour detection operator that is, in general, more effective than that of using push-pull inhibition only. The results of our experiments verify this hypothesis with high statistical significance ($p < 0.001$).

With the rise of deep learning including convolutional neural networks (ConvNets), contour detection has advanced quite substantially in the past few years. The best performing contour detection algorithms that used the Berkeley data set to

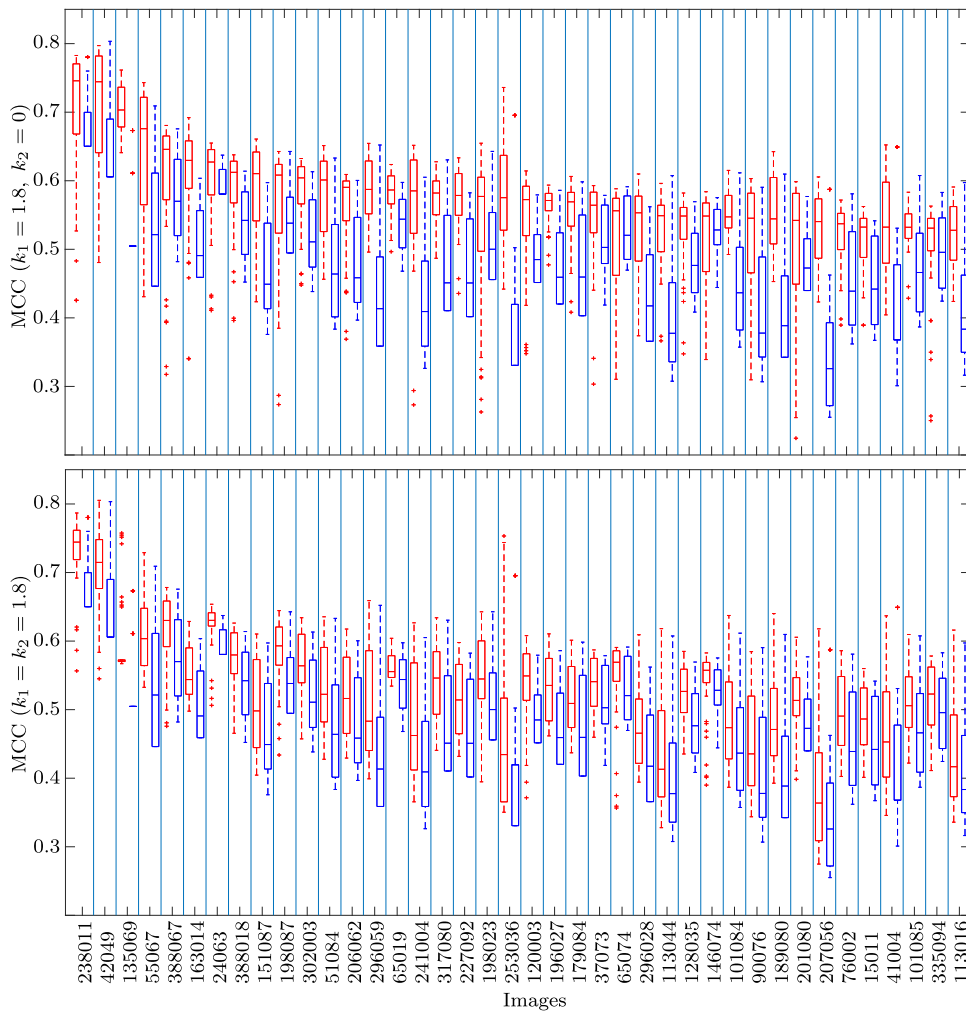


Fig. 8. Box plots of the performance of the new operator against the push-pull CORF model without surround suppression on 40 images from the Berkeley data set. The values of the PP only model (blue) are the same for the top and bottom plots. The red box plots are generated from the results obtained by the proposed model used in two ways: (top) $k_1 = 1.8$ and $k_2 = 0$ and (bottom) $k_1 = k_2 = 1.8$. The labels on the x-axis are the filenames of the respective images. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

evaluate their performance are all based on ConvNets and have been approaching human performance [35–42] with top performers exceeding an F-score of 0.8 for both the Optimal Dataset Scale (ODS) and the Optimal Image Scale (OIS) measurements on the Berkeley test set of 200 images. ConvNets have proven to be a powerful tool in various computer vision applications and they are based on learning hundreds or thousands of linear filters that minimize the error on the training data while preventing over-fitting.

The proposed operator is based on a *single, nonlinear and local* filter. It is inspired by two inhibition phenomena of the visual system of the brain and is aimed to give another point of view on how more effective filters can be designed and applied for contour detection in an unsupervised way. Such an approach belongs to a different research field than that of ConvNets. Our approach belongs to the field of brain-inspired vision operators, which overlaps with computational visual neuroscience, whose aim is to model certain properties of visual neurons. These two approaches, although they can be applied to similar problems, cannot be compared as ConvNets are implemented as full pipelines with abundant learnable feature extractors coupled with a classification model, while the other family of approaches, where ours belongs to, rely on single and local filters. The scope of our work is to demonstrate that the addition of the biological phenomenon of surround suppression to push-pull inhibition results in a much more effective operator than when using only push-pull inhibition. We hope that our results give new insights to the community concerned with ConvNets to investigate the embedding of our novel operator in deep architectures. For instance, in the convolutional layers, next to each simple filter, one may explore adding a pair of a push-pull and a surround suppression filters, which operate within and in the surrounding of the receptive field of the simple filter, respectively. The response maps from the convolutional layers may then be obtained by the function that we propose in this work. We speculate that the introduction of such inhibition filters will improve the generalization

of ConvNets to the different types of noise introduced by the degradation of visual sensors or by changing other conditions (e.g. lighting) that were not represented in the training data. We aim to investigate this approach in future work. In [43], the authors already showed the benefits of adding only the push-pull component in CNN architectures.

In [13], it was shown that a CORF contour detection operator with only push-pull inhibition outperformed various other single-filter-based operators including Canny, Gabor, Gabor with isotropic and anisotropic suppression, and CORF without inhibition. We demonstrate how the new operator that combines push-pull inhibition and surround suppression outperforms (with high statistical significance) the previous CORF operator with push-pull inhibition only. To the best of our knowledge these are the best results ever achieved with a single-filter-based operator on the grayscale converted images of the two concerned benchmark data sets.

As to real-world applications are concerned, many decisions are motivated by cost-benefit analysis. In applications where the collection of training data, though expensive, significantly contributes to an improvement in performance, supervised approaches (e.g. ConvNets) are more appropriate. This is due to their powerful mechanism of learning mapping functions from training data to a desired output. As to generalization, an extensive comparison analysis would be required to investigate the effect of the number of layers and sparsity, for instance, of ConvNets in applications with different conditions and to different types of adversarial attacks. It has already been reported [44] that highly sophisticated ConvNet architectures tend to have less generalization abilities than ones with fewer number of layers. Moreover, real-world vision-based applications, such as visual quality inspection, use visual sensors whose performance may experience degradation over time. Such degradation usually results in different types of noise in the generated images, which may not be immediately visible to the naked eye. ConvNets that are trained on clean images (without noise) may not be robust to images tampered with perturbations [44].

In future work we also aim to investigate the addition of surround suppression in the segmentation of vessel-like structures in 2D and 3D images, such as CT and OCT scans, as well as other applications that rely or use contour detection, such as the ones addressed in [45–50].

As to computational time, on a processor of 2.6GHz Intel Core i7, images of size 321×481 pixels took an average of 0.42s. Since the scope of our work was about the effectiveness of the proposed method rather than its efficiency, the proposed operator was implemented in a sequential model and we did not exploit its highly parallelizable components. In future work, we aim to investigate the best way of implementing this approach on graphical processing units in order to make it appropriate for (near) real-time processing.

5. Conclusions

The contour detection model that we propose is inspired by the phenomena of push-pull inhibition and non-classical receptive fields that occur in many simple cells of visual cortex. The additional computational step of surround suppression in non-classical receptive fields results in a CORF model that is much more effective in suppressing texture, and as a result produces better (with high statistical significance) contour maps of objects within natural scenes.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRedit authorship contribution statement

Damiano Melotti: Software, Validation, Investigation, Writing - original draft, Writing - review & editing. **Kevin Heimbach:** Software, Validation, Investigation, Writing - original draft, Writing - review & editing. **Antonio Rodríguez-Sánchez:** Writing - original draft, Writing - review & editing. **Nicola Strisciuglio:** Writing - original draft, Writing - review & editing. **George Azzopardi:** Conceptualization, Methodology, Investigation, Writing - original draft, Writing - review & editing, Supervision, Project administration.

Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:[10.1016/j.ins.2020.03.026](https://doi.org/10.1016/j.ins.2020.03.026).

References

- [1] D.H. Hubel, T.N. Wiesel, Receptive fields of single neurones in the cat's striate cortex, *J. Physiol.* 148 (3) (1959) 574–591.
- [2] D.H. Hubel, T.N. Wiesel, Receptive fields, binocular interaction and functional architecture in the cat's visual cortex, *J. Physiol.* 160 (1) (1962) 106–154.

- [3] D.H. Hubel, T.N. Wiesel, Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat, *J. Neurophysiol.* 28 (2) (1965) 229–289.
- [4] L.A. Palmer, T.L. Davis, Receptive-field structure in cat striate cortex., *J. Neurophysiol.* 46 (2) (1981) 260–276.
- [5] D. Ferster, Spatially opponent excitation and inhibition in simple cells of the cat visual cortex, *J. Neurosci.* 8 (4) (1988) 1172–1180.
- [6] L.J. Borg-Graham, C. Monier, Y. Fregnac, Visual input evokes transient and strong shunting inhibition in visual cortical neurons, *Nature* 393 (6683) (1998) 369–373.
- [7] J.S. Anderson, M. Carandini, D. Ferster, Orientation tuning of input conductance, excitation, and inhibition in cat primary visual cortex, *J. Neurophysiol.* 84 (2) (2000) 909–926.
- [8] P. Heggelund, Quantitative studies of enhancement and suppression zones in the receptive field of simple cells in cat striate cortex., *J. Physiol.* 373 (1986) 293.
- [9] G.C. DeAngelis, I. Ohzawa, R.D. Freeman, Receptive-field dynamics in the central visual pathways, *Trends Neurosci.* 18 (10) (1995) 451–458.
- [10] N. Petkov, M.A. Westenberg, Suppression of contour perception by band-limited noise and its relation to nonclassical receptive field inhibition, *Biol. Cybern.* 88 (3) (2003) 236–246.
- [11] J.A. Hirsch, L.M. Martinez, Circuits that build visual cortical receptive fields, *Trends Neurosci.* 29 (1) (2006) 30–39.
- [12] D. Ferster, K.D. Miller, Neural mechanisms of orientation selectivity in the visual cortex, *Annu. Rev. Neurosci.* 23 (1) (2000) 441–471.
- [13] G. Azzopardi, A. Rodríguez-Sánchez, J. Piater, N. Petkov, A push-pull corf model of a simple cell with antiphase inhibition improves SNR and contour detection, *PLoS ONE* 9 (7) (2014) e98424.
- [14] D.D. Cox, T. Dean, Neural networks and neuroscience-inspired computer vision, *Curr. Biol.* 24 (18) (2014) R921–R929.
- [15] N. Pinto, Z. Stone, T. Zickler, D. Cox, Scaling up biologically-inspired computer vision: a case study in unconstrained face recognition on facebook, in: *Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2011 IEEE Computer Society Conference on, IEEE, 2011, pp. 35–42.
- [16] T. Serre, L. Wolf, T. Poggio, Object recognition with features inspired by visual cortex, in: *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 2, IEEE, 2005, pp. 994–1000.
- [17] J.G. Daugman, Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters, *JOSA A* 2 (7) (1985) 1160–1169.
- [18] J.P. Jones, L.A. Palmer, An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex, *J. Neurophysiol.* 58 (6) (1987) 1233–1258.
- [19] G. Azzopardi, N. Petkov, A CORF computational model of a simple cell that relies on LGN input outperforms the Gabor function model, *Biol. Cybern.* 106 (3) (2012) 177–189.
- [20] G. Azzopardi, N. Strisciuglio, M. Vento, N. Petkov, Trainable COSFIRE filters for vessel delineation with application to retinal images, *Med. Image Anal.* 19 (1) (2015) 46–57.
- [21] N. Strisciuglio, N. Petkov, Delineation of line patterns in images using B-COSFIRE filters, in: *IWOBI*, 2017, pp. 1–6.
- [22] N. Strisciuglio, G. Azzopardi, N. Petkov, Detection of curved lines with B-COSFIRE filters: a case study on crack delineation, in: *Computer Analysis of Images and Patterns*, 2017, pp. 108–120.
- [23] N. Strisciuglio, G. Azzopardi, N. Petkov, Robust inhibition-augmented operator for delineation of curvilinear structures, *IEEE Transactions on Image Processing* (2019), 1–1
- [24] P. Bishop, J.S. Coombs, G. Henry, Receptive fields of simple cells in the cat striate cortex, *J. Physiol.* 231 (1) (1973) 31.
- [25] D.J. Field, A. Hayes, R.F. Hess, Contour integration by the human visual system: evidence for a local association field, *Vis. Res.* 33 (2) (1993) 173–193.
- [26] R.N. Sachdev, M.R. Krause, J.A. Mazer, Surround suppression and sparse coding in visual and barrel cortices, *Front Neural Circuits* 6 (2012) 43.
- [27] C. Grigorescu, N. Petkov, M.A. Westenberg, Contour detection based on nonclassical receptive field inhibition, *IEEE Trans. Image Process.* 12 (7) (2003) 729–739.
- [28] C. Grigorescu, N. Petkov, M.A. Westenberg, Contour and boundary detection improved by surround suppression of texture edges, *Image Vis. Comput.* 22 (8) (2004) 609–622.
- [29] C. Zeng, Y. Li, K. Yang, C. Li, Contour detection based on a non-classical receptive field model with butterfly-shaped inhibition subregions, *Neurocomputing* 74 (10) (2011) 1527–1534.
- [30] B.-h. Liu, Y.-t. Li, W.-p. Ma, C.-j. Pan, L.I. Zhang, H.W. Tao, Broad inhibition sharpens orientation selectivity by expanding input dynamic range in mouse simple cells, *Neuron* 71 (3) (2011) 542–554.
- [31] Y.-t. Li, W.-p. Ma, L.-y. Li, A. Ibrahim, S.-z. Wang, H.W. Tao, Broadening of inhibitory tuning underlies contrast-dependent sharpening of orientation selectivity in mouse visual cortex, *J. Neurosci.* 32 (46) (2012) 16466–16477.
- [32] J. Canny, A computational approach to edge detection, *IEEE Trans. Pattern Anal. Mach. Intell.* (6) (1986) 679–698.
- [33] F. Arbelaez Maire, Malik, Contour detection and hierarchical image segmentation, *IEEE Trans. Pattern Anal. Mach. Intell.* 33 (5) (2011) 898–916.
- [34] D.R. Martin, C.C. Fowlkes, J. Malik, Learning to detect natural image boundaries using local brightness, color, and texture cues, *IEEE Trans. Pattern Anal. Mach. Intell.* 26 (5) (2004) 530–549.
- [35] Q. Hou, J. Liu, M. Cheng, A. Borji, P.H.S. Torr, Three birds one stone: a unified framework for salient object segmentation, edge detection and skeleton extraction, *arXiv:1803.09860* (2018).
- [36] Y. Liu, M. Cheng, X. Hu, K. Wang, X. Bai, Richer convolutional features for edge detection, *arXiv:1612.02103* (2016).
- [37] I. Kokkinos, Pushing the boundaries of boundary detection using deep learning, 2016.
- [38] Y. Wang, X. Zhao, Y. Li, K. Huang, Deep crisp boundaries: from boundaries to higher-level tasks, *arXiv:1801.02439* (2018).
- [39] Y. Liu, M.S. Lew, Learning relaxed deep supervision for better edge detection, in: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 231–240.
- [40] J. Yang, B.L. Price, S. Cohen, H. Lee, M. Yang, Object contour detection with a fully convolutional encoder-decoder network, *arXiv:1603.04530* (2016).
- [41] Wei Shen, Xinggang Wang, Yan Wang, Xiang Bai, Z. Zhang, DeepContour: a deep convolutional feature learned by positive-sharing loss for contour detection, in: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 3982–3991.
- [42] G. Bertasius, J. Shi, L. Torresani, DeepEdge: a multi-scale bifurcated deep network for top-down contour detection, 2015, pp. 4380–4389.
- [43] N. Strisciuglio, M. Lopez-Antequera, N. Petkov, Enhanced robustness of convolutional networks with a push pull inhibition layer, *Neural Comput. Appl.* (2020), doi:10.1007/s00521-020-04751-8.
- [44] D. Hendrycks, T.G. Dietterich, Benchmarking neural network robustness to common corruptions and perturbations, *arXiv:1903.12261* (2019).
- [45] X. Lin, Z.-J. Wang, X. Tan, M.-E. Fang, N.N. Xiong, L. Ma, MCCH: a novel convex hull prior based solution for saliency detection, *Inf. Sci.* 485 (2019) 521–539.
- [46] Z. Chen, R. Wang, Z. Zhang, H. Wang, L. Xu, Background foreground interaction for moving object detection in dynamic scenes, *Inf. Sci.* 483 (2019) 65–81.
- [47] J. Yu, B. Zhang, Z. Kuang, D. Lin, J. Fan, Iprivacy: image privacy protection by identifying sensitive objects via deep multi-task learning, *IEEE Trans. Inf. Forensics Secur.* 12 (5) (2017) 1005–1016.
- [48] J. Zhang, J. Yu, D. Tao, Local deep-feature alignment for unsupervised dimension reduction, *IEEE Trans. Image Process.* 27 (5) (2018) 2420–2432.
- [49] J. Yu, C. Zhu, J. Zhang, Q. Huang, D. Tao, Spatial pyramid-enhanced NetVLAD with weighted triplet loss for place recognition, *IEEE Trans. Neural Netw. Learn. Syst.* (2019) 1–14.
- [50] J. Yu, M. Tan, H. Zhang, D. Tao, Y. Rui, Hierarchical deep click feature prediction for fine-grained image recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2019), 1–1