

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Surround Suppression and Recurrent Interactions V1-V2 for Natural Scene Boundary Detection

Francisco J. Díaz-Pernas, Míriam Antón-Rodríguez, Isabel de la Torre-Díez,
Mario Martínez-Zarzuela, David González-Ortega, Daniel Boto-Giralda,
and J. Fernando Díez-Higuera

*Department of Signal Theory, Communications and Telematics Engineering
Telecommunications Engineering School, Valladolid University, Valladolid
Spain*

1. Introduction

Understanding the brain functioning in visual process is one of the more active areas within neuroscience, modeling and computation. This great interest is due to the fact that vision provides us with the more important information about the surroundings. Oriented early filtering processes, perceptual clustering through emerging features, form and depth perception, figure-ground separation, and object recognition are all involved in the brain visual activity. Important researches have been undertaken to develop models simulating this brain behavior. So, neural models of visual perception biologically motivated in early vision tasks have been arisen versus others not biologically motivated computer vision algorithms. The knowledge about the early vision is notable and, accordingly, the neural modeling of the early vision tasks has been very scientifically productive. There are numerous visual models of the early visual perception (Kokkinos et al., 2007) (Neumann et al., 2007). One of the more significant is the BCS/FCS model of Grossberg and Mingolla (Grossberg & Mingolla, 1985). This model is composed of two systems, boundary contour system (BCS) and feature contour system (FCS), modeling processes given in the interactions among V1, V2 and V4 visual areas. Features so important in the human vision like illusory perception, emergent segmentations, diffusive filling-in are integrated in a coherent way in the BCS/FCS network. This system has experiment an important evolution in its modeling, integrating spatial and orientational competitive processes in a same stage (Mingolla et al., 1999). This model has been the framework of many other researchers in the development of their approaches. Kokkinos et al. (Kokkinos et al., 2008) have recently developed a computational simplification of the BCS/FCS model. They propound a three-stage structure incorporating feature extraction with contrast normalization, boundary formation through a competitive stage of information of feature extraction, cooperative signal, anisotropic image smoothing and large scale signal. In their proposal, they include a comparison to Canny's detector, a classical computer vision algorithm. Kokkinos' architecture shows better results than Canny's algorithm. Main contribution of Kokkinos' is its simplification of the BCS/FCS model more than its evolution as it is somewhat complex.

Several neurophysiological studies show that the response modulation by stimuli beyond the receptive field is a common phenomenon in the V1 area (Petkov & Westenberg, 2003). There are evidences that 80% of the orientation selective cells perform a response suppression over a bar texture surround beyond its receptive field. Based on these studies Petkov and Westenberg (Petkov & Westenberg, 2003) proposed a surround suppression model for the complex cells and they also analyzed this effect by psychophysical experiments. This interesting inhibitory effect has been considered in the development of our proposal of boundary detection neural network. This surround suppression mechanism is not considered in the BCS. Nevertheless, Petrov and Westenberg's network is limited to the simple and complex cells for evaluating the surround suppression effect. Competitive and cooperative processes generated by the V2 hypercomplex cells, are not considered in their proposal. Another important biological evidence is the existence of feedback interactions in the visual cortex (Hubel, 1995). BCS includes a competitive-cooperative loop where the completion boundary and the illusory contour extraction are performed. Neumann and Sepp (Neumann & Sepp, 1999) presented a V1-V2 recurrent process through the integration of the V1 activations from V2 receptive field elongated lobes.

In the work presented in this paper, trying to obtain a model as simple as possible, we propose an architecture considering the more significant mechanisms of the human visual system early processing for color scene stimuli: chromatic and achromatic opponent channels, orientational filtering, surround suppression, V1-V2 recurrent interactions through competitive and cooperative fields. Additionally, we propose an inter-scale information fusion stage in order to obtain the boundary output with all the information gathered from the scene. In (Antón-Rodríguez et al., 2009) we proposed a visual feature extraction architecture for color-texture identification corresponding to the color extension of the BCS/FCS system. In this work we presented a BCS system for processing signals from three channels, two opponent chromatic channels and a luminance one, and two FCS for diffusing the two chromatic channels.

In the present work, we propose a new model with six opponent channels, four chromatic and two achromatic, emerged from the transformation of the RGB image. It also includes new mechanisms, like surround suppression and inter-scale fusion, to achieve the natural scene boundaries.

2. Perceptual boundary recurrent detection neural architecture

The Perceptual boundaRy rEcurrent dEtectiOn Neural (PREEN) proposed model (see Fig. 1) comprises five main components, respectively designated as Colour Opponent stage (CO), Chromatic Contour stage (CC), Competitive Fusion stage (CF), Contour Saliency stage (CS), and Inter-scale Competition (IC). The neural model processing is achieved through multiple spatial scales.

The CO stage transforms the chromatic components of the input signals (RGB) into a bio-inspired codification system, made up of various opponent chromatic channels and an achromatic channel. In order to do this, the CO stage firstly calculates the activations of the long- (L), middle- (M), and short- (S) wavelength retinal cones, and then, it generates the opponent processes, corresponding to the ON-OFF achromatic channel and the $L+M$ -, $M+L$ -, $S+(L+M)$ - and $(L+M)+S$ - chromatic opponencies. Studies of the human visual system have found that visual stimuli take part in the color opponent and enhancement processes located in retina and Lateral Geniculate Nucleus (LGN) cells of the mammalian visual system (Hubel, 1995; Wilson et al., 1990).

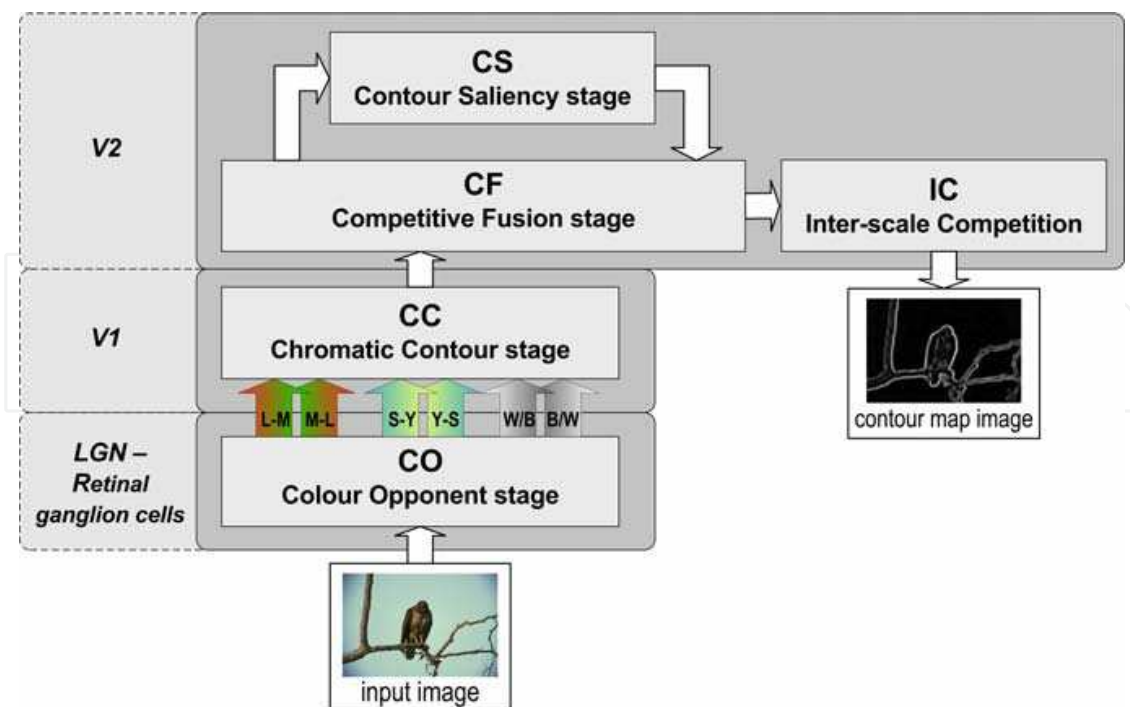


Fig. 1. Proposed model architecture

CC stage is composed of simple and complex cells which filter and perform a surround suppression through multiple scales and orientations in order to extract the real boundaries of the scene. Simple cell multi-scale filtering extracts textural features from both the color opponent and the achromatic signals.

Complex cells use a competition network with surround suppression, which attenuates the contour activities from inner texture areas where it appears a high concentration of false boundaries. Complex cells coherently fuse the color and texture activities, generated by the simple cells, with a behavior inspired in cells from the V1 visual area of the human visual system (Petrov & Westenberg, 2003).

Biologically inspired V1-V2 recurrent interactions take place in the PREEN architecture. CF stage competitively fuses top-down and bottom-up signals, complex-CF-CS, defining and shaping the natural boundaries. CF and CS stages model competitive and cooperative behaviors exhibited in the V2 area of the human visual system (Hubel, 1995).

Final output is constituted by a contour map image corresponding to the natural boundaries with perceptual significance.

The PREEN model architecture includes competitive networks. We model these networks using the membrane potential network (Hodgkin & Huxley, 1990), whose simplified behavior can be expressed according to equation (1).

$$C_M \frac{dV}{dt} = -AV + (V^+ - V)g^+ + (V^- - V)g^- \tag{1}$$

where V is the membrane potential, A is a decay constant, V^+ and V^- are the excitatory and inhibitory reversal potentials, and g^+ and g^- are the excitatory and inhibitory total input, respectively.

In a stationary situation, the potential, V , would be defined by equation (2). This is the situation we consider in our model.

$$V = \frac{V^+ g^+ + V^- g^-}{A + g^+ + g^-} \quad (2)$$

where $V^+ \geq 0$ and $V^- \leq 0$.

Equation (2) expresses the membrane potential as the normalization between the net input (difference between excitation and inhibition) and the total input (excitation plus inhibition). So, this normalization computes the ratio contrast and solves the noise-saturation dilemma.

2.1 Colour opponent stage

The CO stage performs chromatic and achromatic opponent competitive processes based on opponent mechanisms and luminance channel competitive enhancement to generate four chromatic signals shaping opponent pairs, L^+M^- , M^+L^- , $S^+(L+M)^-$ and $(L+M)^+S^-$, and two achromatic signals, ON and OFF. This is observed as the generation of three channels of opponent pairs. In hierarchical levels of the human system there are evidences of manifold cases of opponent pairs, both in the motor and visual systems (Zrenner, 1990).

The CO processing contains a previous stage in which luminance (I signal), activations of the long (L signal), middle (M signal), short (S signal) wavelength cones and (L+M) channel activation (Y signal) are generated from R, G and B input signals.

Equations (3), (4) and (5) define the calculations of those activities (Antón-Rodríguez et al., 2009). The luminance signal (I) is computed as a weighted sum; the L, M and S signals are obtained as the transformation matrix of the three chromatic components R, G and B.

$$I = 0.299R + 0.587G + 0.114B \quad (3)$$

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.293 & 0.603 & 0.104 \\ 0.134 & 0.704 & 0.162 \\ 0.046 & 0.099 & 0.854 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (4)$$

$$Y = L + M \quad (5)$$

The CO stage models the behavior of the opponent cells whose main characteristics are their high sensibility and precision to contrasts contours. These attributes recommend their use as a previous step for image contour, shapes, and texture detection (Hubel, 1995). The opponent model for L^+M^- chromatic channels, following membrane competitive network is defined by equation (6).

$$c_{ij}^{lm(s)} = \frac{\sum [D_{pq}^{(s)}]^+ L_{ij} - \sum [-D_{pq}^{(s)}]^+ M_{ij}}{A + \sum [D_{pq}^{(s)}]^+ L_{ij} + \sum [-D_{pq}^{(s)}]^+ M_{ij}} \quad (6)$$

where $c_{ij}^{lm(s)}$ is the L^+M^- channel cell activity for position (i,j) , $[c]^+ = \max(0, c)$, A is a decay constant, and $D_{pq}^{(s)}$ is a difference of Gaussians (see Fig. 2) following equation (7).

$$D_{pq}^{(s)} = D_A \left[\exp\left(\frac{-(p^2 + q^2)}{2\sigma_e^{(s)2}}\right) - \exp\left(\frac{-(p^2 + q^2)}{2\sigma_i^{(s)2}}\right) \right] \quad (7)$$

where σ_e and σ_i are the excitation and inhibition deviations and D_A is the amplitude constant.

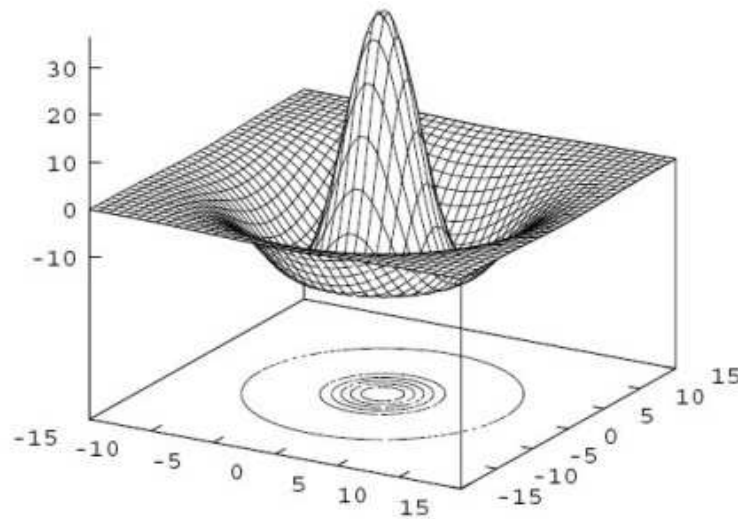


Fig. 2. Difference of Gaussians (dog)

The remainder opponencies, $M+L-$, $S+Y-$ and $Y+S-$, are calculated using equation (6) substituting L_{ij} and M_{ij} signals for the corresponding opponency signals.

Analogously to the achromatic process accomplished in the LGN of the human visual system (Hubel, 1995), the neural architecture enhances the luminance signal, I_{ij} , in a competitive network generating two opponent channels ON-center and OFF-center under equation (6), similarly to the chromatic channels, replacing the chromatic signals by the luminance signal. ON channel processing is in charge of enhancing the information of regions with higher luminance intensity in relation to their surroundings, while OFF channel enhances regions with lower luminance intensity than their surroundings.

These competitive processes, derived from equation (6), establish a gain control network over the inputs from chromatic and luminance channels, maintaining the sensibility of cells to contrasts, compensating variable illumination, and normalizing image intensity.

These competitive networks are parallelly applied over three different spatial scales, with suitable values of σ_e and σ_i for the small, medium and large scales ($s=0,1,2$).

Correlation studies about color codification systems based on trichromatic theory have proven the suitability of using the red-green and blue-yellow opponent signals (Ohta et al., 1980), in all, opponencies of long-middle and short-long-middle wavelength. In like manner, the ganglionar opponent cells with higher presence rate in the retina correspond to L-M opponency, next higher rate is S-(L+M) (Zrenner et al., 1990), so we consider appropriate using the opponencies chosen.

Biologically motivated models for boundary detection, as BCS/FCS (Grossberg et al., 1995) or Kokkinos' model (Kokkinos et al., 2008) do not use chromatic channels for detecting boundaries. They only work with the luminance component. Nevertheless, they use images of natural scenes for validation. It is clearly that the human visual system operating has an essential chromatic component, based on the color opponencies.

2.2 Chromatic Contour (CC)

The Chromatic Contour stage models V1 simple and complex cells. It undertakes the contour extraction and the merging with surround suppression of the simple cell signals, as it can be seen in Fig. 1. CC stage models behaviors of cells located in V1 visual area, orientation and spatial frequency selective cells and information fusion cells (Hubel, 1995). Simple cells extract contours from chromatic and achromatic CO channels through a Gabor filter bank, using even ($E_{ijk}^{(s)}$) and odd ($O_{ijk}^{(s)}$) components for position (i,j), deviation for multiple scales ($s=0, 1, 2$) and orientation ($k=0, 1, 2, 3, 4, 5$ corresponding to $\theta=30^\circ, 60^\circ, 90^\circ, 120^\circ, 150^\circ$). These cells respond to variations on the textural and color features. Complex Gabor filters have sensibility to orientation, spatial frequency and position (Daugman, 1980). Grossberg, Mingolla et al. (Grossberg et al., 1995) (Mingolla et al., 1999) modeled simple cells with an odd-type filtering in their BCS model. The inclusion of even fields is justified due to the importance of these profiles in texture detection (Landy & Bergen, 1991). Following the behavior indicated by the membrane potential shown in equation (2), simple cell activities are given by equations (8), (9), (10) and (11).

$$a_{ijk}^{(s)} = \frac{\sum E_{pqk}^{(s)} \left(\left[c_{ij}^{on(s)} \right]^+ - \left[c_{ij}^{off(s)} \right]^+ \right)}{A + \sum |E_{pqk}^{(s)}| \left(\left[c_{ij}^{on(s)} \right]^+ + \left[c_{ij}^{off(s)} \right]^+ \right)} \quad (8)$$

$$b_{ijk}^{(s)} = \frac{\sum O_{pqk}^{(s)} \left(\left[c_{ij}^{on(s)} \right]^+ - \left[c_{ij}^{off(s)} \right]^+ \right)}{A + \sum |O_{pqk}^{(s)}| \left(\left[c_{ij}^{on(s)} \right]^+ + \left[c_{ij}^{off(s)} \right]^+ \right)} \quad (9)$$

$$e_{ijk}^{lm(s)} = \frac{\sum E_{pqk}^{(s)} c_{ij}^{lm(s)}}{A + \sum |E_{pqk}^{(s)}| c_{ij}^{lm(s)}} \quad (10)$$

$$f_{ijk}^{lm(s)} = \frac{\sum O_{pqk}^{(s)} c_{ij}^{lm(s)}}{A + \sum |O_{pqk}^{(s)}| c_{ij}^{lm(s)}} \quad (11)$$

where $a_{ijk}^{(s)}$, $b_{ijk}^{(s)}$ are the even and odd simple cell activities for the luminance channel, and $e_{ijk}^{lm(s)}$, $f_{ijk}^{lm(s)}$ are the simple cell activities for chromatic channel L+M-, position (i,j) and orientation k for even and odd filters respectively, $|\cdot|$ represents the absolute value, $[c]^+ = \max(0, c)$, and A is a decay constant.

For the remainder chromatic channels, the activity equations to use are equations (10) and (11) replacing L+M- channel signal by the appropriate channel signal (M+L-, S+Y-, Y+S-).

For each perceptual position (i,j) a hyper column of simple cells varying in filter type (filter profile) and orientation k is applied to each channel.

The complex cell stage, using two cellular layers, fuses information from simple cells giving rise to a map which contains real contours for each of the three scales used (see Fig. 1).

The first layer of cells is in charge of combining responses from different Gabor filters at each opponent and luminance channel at their three scales ($s=0,1,2$), as shown in equation (12).

$$h_{ijk}^{(s)} = \left[a_{ijk}^{(s)} \right]^+ + \left[b_{ijk}^{(s)} \right]^+ + \left[e_{ijk}^{lm(s)} \right]^+ + \left[f_{ijk}^{lm(s)} \right]^+ + \left[e_{ijk}^{slm(s)} \right]^+ + \left[f_{ijk}^{slm(s)} \right]^+ + \left[e_{ijk}^{ml(s)} \right]^+ + \left[f_{ijk}^{ml(s)} \right]^+ + \left[e_{ijk}^{lms(s)} \right]^+ + \left[f_{ijk}^{lms(s)} \right]^+ \quad (12)$$

where $[c]^+ = \max(0, c)$ is a half-wave rectifier.

The second complex cell layer models a competition network with surround suppression (see Fig. 3).

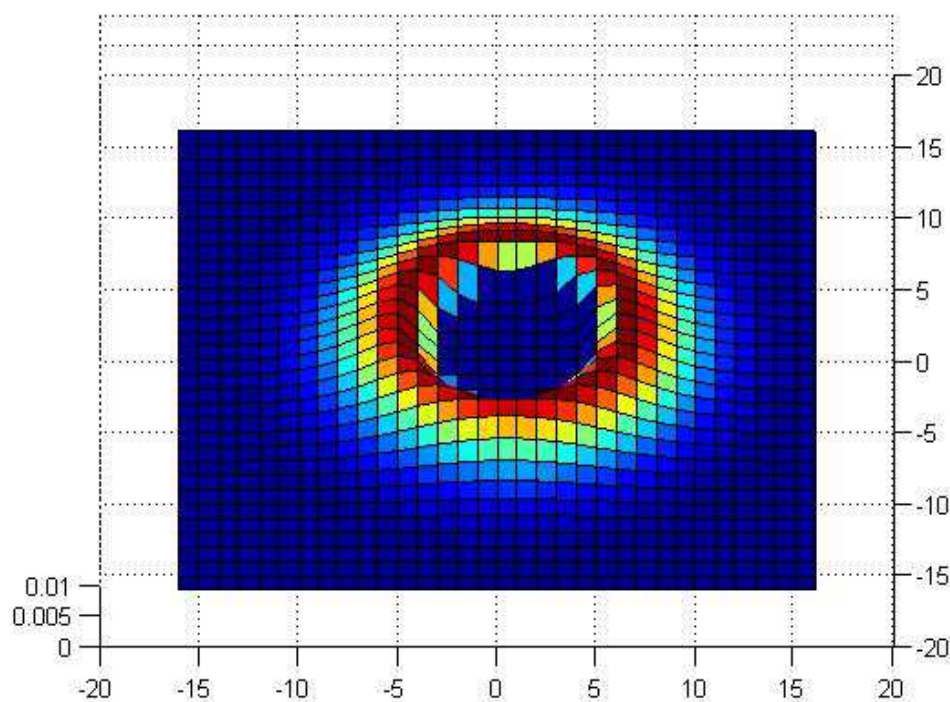


Fig. 3. Surround suppression receptive field of CC complex cells ($S_{ijk}^{(s)}$).

The behavior of layer 2 complex cells is modeled by equation (13):

$$H_{ijk}^{(s)} = \left[\frac{h_{ijk}^{(s)} - \lambda \sum_{pqk} S_{pqk}^{(s)} h_{ij}^{(s)}}{A + h_{ijk}^{(s)} + \sum_{pqk} S_{pqk}^{(s)} h_{ij}^{lm(s)}} \right]^+ \quad (13)$$

being $H_{ijk}^{(s)}$ the CC complex cells output; $h_{ijk}^{(s)}$ is the first layer activity, $S_{ijk}^{(s)} = [-D_{ij}^{(s)}]^+$ is the surround suppression receptive field with $D_{ij}^{(s)}$ a dog following equation (7), as it is shown in Fig. 3, for each of the scales defined by $s=0,1,2$; A is a decay constant, $\lambda > 0$ is the suppression constant and $[c]^+ = \max(0, c)$. Equation (13) determines the anisotropic surround suppression, as the removal is performed according to the orientations.

In their work, Petrov and Westenberg (Petrov & Westenberg, 2003) observed various manifestations of perceptual modulation by the context. They used image sets of letters and bars, and objects icons (of different sizes and orientations) with superimposed band-spectrum noise. They observed the effects of the surround in contrast for different noise frequencies. The inhibitory effect was modeled designing a non-classical receptive field (non-CRF) defined by a weighting function with a normalized difference of Gaussians profile. This idea is taken in the present paper but with an important difference.

Model proposed of CC complex cells exhibits a positional competition between a position (i,j) and its surroundings. These competitive processes establish a gain control network over the fusion channel input, maintaining the sensibility of cells to contrasts, compensating variable illumination, and normalizing image intensity.

In order to obtain a complex non-CRF field external to the CRF field of CC simple cells, we choose same values for σ_e deviations of $S_{ijk}^{(s)}$ and for $\sigma^{(s)}$ deviations of CC simple cells and we fix σ_i deviations of $S_{ijk}^{(s)}$ as twice the value of $\sigma^{(s)}$.

In Fig. 4 it can be observed the effect produced by the surround suppression depending on the inhibition strength ($\lambda=1.0$ and $\lambda=2.0$). Using a high suppression constant, it is possible to remove areas with a high concentration of weak contour caused by inner texture features (see left column). It can be noticed that the surround suppression model makes the boundaries to have less noisy activities near them, thus achieving a better definition and so a better precision. However, as the surround suppression model cleans up the area around the boundaries, sometimes the recall value diminishes.

In a later work, Grigorescu, Petrov and Westenberg (Grigorescu et al., 2004) proposed a computational step, called surround suppression for detecting boundaries in natural scenes. This step was incorporated to the Canny edge detector. A comparison was performed using 40 natural images, achieving better results with the step included.

To strengthen the analysis about the positive effect of the surround suppression modeling, we took 20 images from the Berkeley Segmentation DataSet (BSDS) and compared the results obtained with and without suppression. In Fig. 4, we can see some processing examples. Fig. 4 (e) includes the F-value curve of the processing with (right) and without suppression (left). We achieved a mean F-value of 0.64 (0.62, 0.66) when processing with suppression versus a mean value of 0.63 (0.69, 0.57) when processing without suppression. It can be observed a slightly better result counting the suppression. A remarkable feature is the significant difference in the precision value, in favor of the model with suppression, and in the recall value, in favor of the model without suppression, which corroborates the previous point.

Using cooperation processes, it is more interesting obtaining high precision than high recall, since these processes generate more boundaries, but also more noise. Accordingly, the model of complex cells showing surround suppression advantages the cooperative processes from V2, which PREEN model includes in the CS stage.

The explained complex cell stage modeling has significant differences with other processing models. The complex cell stage of the BCS model (Grossberg et al., 1995) is in charge of summing simple cell signals. This sum produces an independence of the contrast direction, so as the entire real boundary map is obtained in this stage.

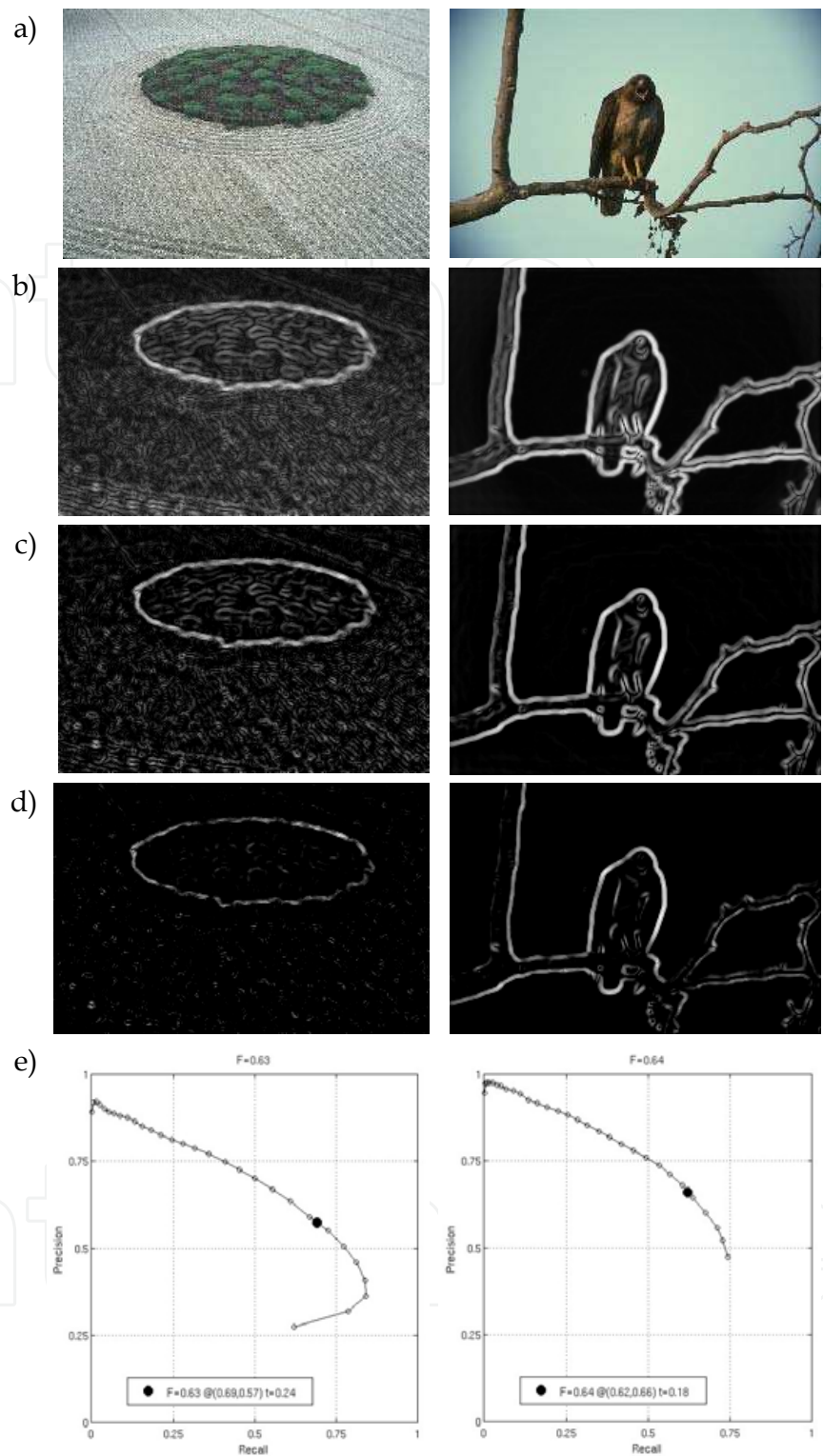


Fig. 4. Effect of complex cell surround suppression. Image on the left: image 86016. Image on the right: image 42049. Row a): original image. Row b): output from first stage (fusion) of complex stage (without surround suppression). Row c): output from CC complex stage for $\lambda=1.0$. Row d): output from CC complex for $\lambda=2.0$. Row e): F-measure curves (precision and recall) for 20 test images of the Berkeley database, (e)-left without surround suppression and (e)-right with surround suppression.

In PREEN model, the complex cell stage fuses simple cell signals and also removes the surroundings making boundaries to be contrast-independent and more accurate.

Kokkinos et al.'s model uses a shunting inhibition for surround suppression which implies a recurrent stage with obliged stabilization (Kokkinos et al., 2008). PREEN model uses a non-recurrent competition, as the recurrent process is located in upper stages, where V1-V2 interaction occurs.

2.3 Competitive Fusion stage (CF)

CF stage competitively fuses information from CC complex bottom-up and CS cooperation top-down. This stage models the V1-V2 recurrent interactions produced in the PREEN model. This recurrent interaction detects, regulates, and completes boundaries into globally consistent contrast positions and orientations, while it suppresses activations from redundant and less important contours, thus eliminating image noise. There are numerous evidences about the existence of hyper-complex cells in the human visual system, where V1-V2 recurrent interactions occurs (Hubel, 1995). We use a shunting network to model this behavior, where lateral inhibitions occur within cells of the same competitive stage. Equation (14) describes these CF hyper-complex cells functioning. It includes two inhibitions, orientationally and spatially.

$$U_{ijk}^{(s)} = \left[\frac{u_{ijk}^{(s)} - C_i \sum_{r \neq k} \sum_{pq} S_{pqr}^{(s)} U_{pqr}^{(s)} - C_c \sum_{pq} G_{pq}^{(s)} U_{pqk}^{(s)}}{A + u_{ijk}^{(s)} + \sum_{pq} G_{pq}^{(s)} U_{pqk}^{(s)}} \right]^+ \quad (14)$$

being $U_{ijk}^{(s)}$ the CF hyper-complex cells output, $u_{ijk}^{(s)} = K_h H_{ijk}^{(s)} + K_f F_{ijk}^{(s)}$ the input to the competition stage: $H_{ijk}^{(s)}$ is the CC complex cells output, $F_{ijk}^{(s)}$ is the CS top-down signal, K_h and K_f are gain constants; the orientational inhibitory receptive field, $S_{pqr}^{(s)}$, is a Gaussian kernel with τ ($\tau > 1$) aspect ratio, rotated r grades, and with $\sigma_i^{(s)}$ deviation for each of the scales defined by $s=0,1,2$; $G_{pq}^{(s)}$ is a Gaussian kernel with $\sigma_u^{(s)}$ deviation for each of the scales, A is a decay constant, C_c and C_i are inhibition constants and $[c]^+ = \max(0, c)$.

Recurrent interaction is solved by an iterative process, where actual state of the $U_{ijk}^{(s)}$ activity depends on the previous state activity and the CS stage feedback activity, $F_{ijk}^{(s)}$.

The inhibitory receptive field in the competition among orientations, $S_{pqr}^{(s)}$, corresponds to a Gaussian rotated r grades, with a very high aspect ratio so that all the points in the line of orientation r passing through (i,j) have a maximum inhibition (see Fig. 5).

The point of difference between competitive stages of PREEN model and BCS system is the competitive model with shunting inhibition used in PREEN model versus a contrast normalization network from BCS system. Shunting inhibition allows higher boundary enhancement and thinning yielding higher precision and deletion of noise activities, which work against cooperative mechanisms, advantaging the recurrent process.

Kokkinos’ architecture also uses a shunting inhibition model, including in the competition a gradient signal from the smooth stage through a diffusion process, similar to the FCS stage of the BCS/FCS system. Kokkinos’ model does not perform a competition among orientations within this stage. PREEN architecture, similarly to BCS system, proposes a competition among orientations to better define the boundary orientation in each position. In this competition, PREEN uses the inhibition signal from competitive stage neighbor instead of the input signal to the competitive stage taken in the BCS model. That is to say PREEN model heightens the shunting mechanism in the inhibition.

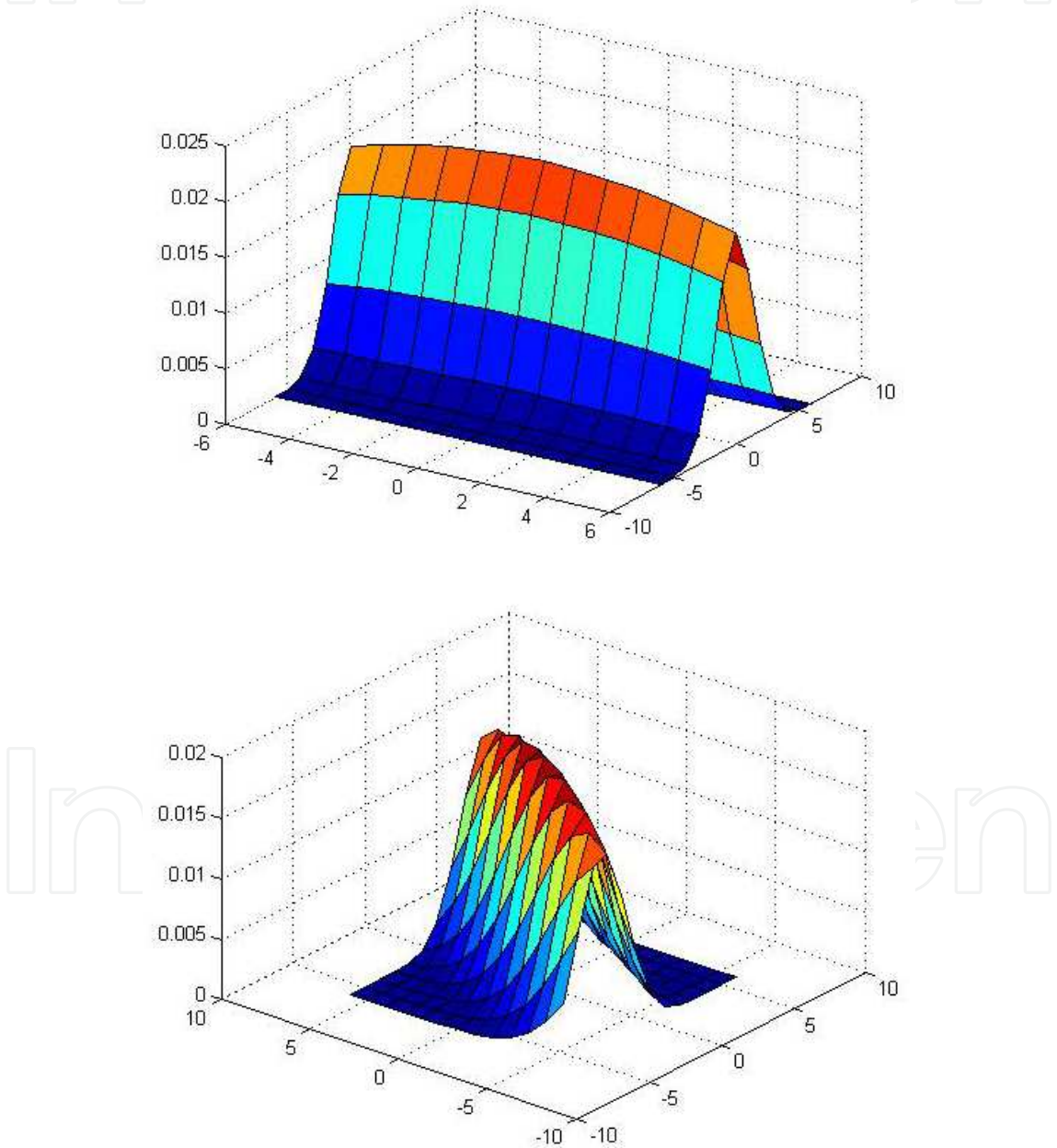


Fig. 5. Orientational inhibition receptive field, $S_{pqr}^{(s)}$.

2.4 Contour saliency stage (CS)

CS stage models the cooperative process for the saliency detection of the contour activity interpolation. In visual cortex, it has been detected cooperative completion or grouping processes by realizing a “bipole property” (Mingolla et al., 1999). This property refers to the disposition of certain cells in visual cortical areas V1 and V2 to fire in the presence of approximately aligned, but spatially separated, image gradients. Salient contours are detected determining existent contour activity in both sides of each position (i,j) through a bipole receptive field. The behavior of CS cells is modeled by equation (15).

$$F_{ijk}^{(s)} = \frac{z\left(\sum P_{pqk}^{(s)} U_{pqk}^{(s)}\right) z\left(\sum N_{pqk}^{(s)} U_{pqk}^{(s)}\right)}{A + \left(\sum P_{pqk}^{(s)} U_{pqk}^{(s)}\right) \left(\sum N_{pqk}^{(s)} U_{pqk}^{(s)}\right)} \quad (15)$$

where $U_{ijk}^{(s)}$ is the CF hyper-complex activity, $P_{ijk}^{(s)}$ and $N_{ijk}^{(s)}$ are the lobes of the bipole receptive field of CS cells (see Fig. 6), A is a decay constant, and $z(s)$ is a lineal function with a positive cooperative threshold, α :

$$z(s) = \begin{cases} 0 & \text{si } s \leq \alpha \\ s & \text{si } s > \alpha \end{cases}$$

Each lobe is generated as a Gaussian with an aspect ratio of τ and a profile relation of $K_l^{(s)}$ for each scale s , with a coordinate translation D_c and a rotation k , following equation (16). Fig. 6 shows a dipole example.

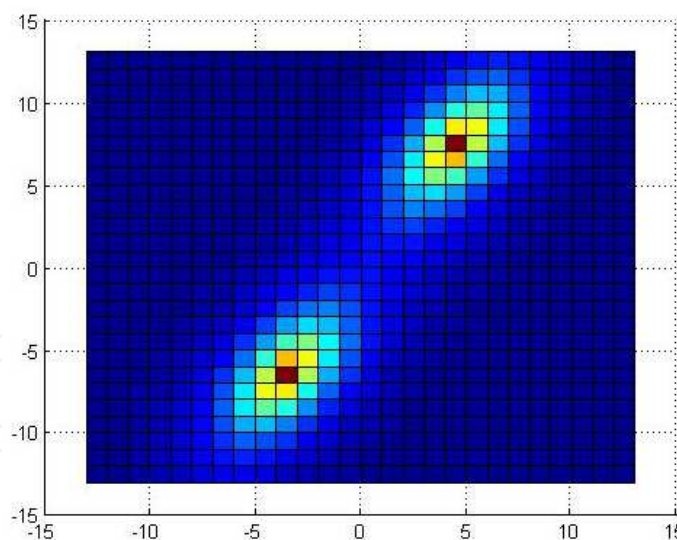


Fig. 6. Oriented bipole receptive field.

$$BIP_{ijk}^{(s)} = \exp \left(-K_l^{(s)} \sqrt{\left(\frac{x'}{\tau}\right)^2 + (y')^2} \right) \quad (16)$$

$$P_{ijk}^{(s)} = \frac{1}{\|BIP_{ijk}^{(s)}\|} BIP_{ijk}^{(s)}$$

where $\|\cdot\|$ is the L1 norm, the coordinates of the spatial transformation:

$$\begin{aligned}x' &= (x - x_c) \cos(k) + (y - y_c) \sin(k) \\y' &= -(x - x_c) \sin(k) + (y - y_c) \cos(k)\end{aligned}$$

and the translation origin coordinates are:

$$\begin{aligned}x_c &= D_c \cos(k) \\y_c &= D_c \sin(k)\end{aligned}$$

Equations for computing $N_{ijk}^{(s)}$ are similar but with a coordinate translation of $-D_c$. Each lobe measures one unit of area, and it is essential to be no activity in the dipole center. Competitive stage output $U_{ijk}^{(s)}$ is normalized, bounded between 0 and 1. Hence, maximum activity in each lobule will be 1. Thereby, we can appropriately choose the value of α . Thanks to this contour cooperative saliency, illusory boundaries of perceptual figures are extracted as it can be observed in Fig. 7, where it appears the PREEN model processing result of the Kanizsa's square image.

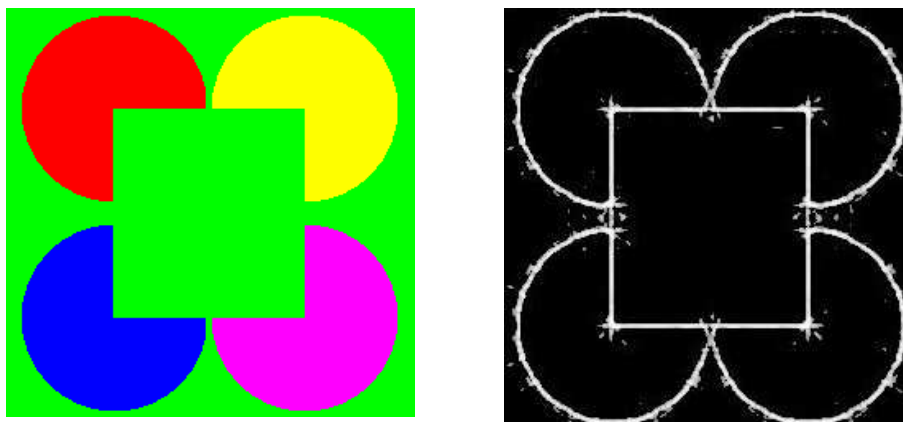


Fig. 7. PREEN processing of Kanizsa's square image.

In BCS system, Kokkinos' model, and PREEN architecture cooperation is performed in a similar way. All uses bipole receptive field cells to generate the cooperative activity produced when both lobes of the receptive field are active. Difference between both models lie in the way of determining when such excitatory situation is produced. As it is expressed in equation (15), PREEN uses the product of the lobe activities to obtain the cooperation activity. Each lobe contributes to excite the dipole cell is its activity exceeds the threshold α . Hence, the product demands the contribution of both lobes and, so, a normalized bipole cell output near 1 will be guaranteed. In remainder cases, the output will be 0.

Kokkinos utilizes the product of the lobule activities but without activation threshold. And BCS uses a pseudo-logic function of decision AND, with a fast saturation function to a value near to 1, and with a threshold of 2 (Mingolla et al., 1999). By doing so, it is guaranteed that the bipole cell exceed 2 only if there are inputs to both lobules.

2.5 Inter-scale Competition stage (IC)

PREEN model proposes an output stage where signals from the three scales are merged in order to obtain the best possible output combining all the information gathered in each scale. While scale increases, noise is better filtered and so boundary signals improve, though precision and thinning become reduced. Noise does not appear in larger scales, but only contours with higher gradient are detected; those are high quality contours but not all existing. Boundaries from the small scale are accurate and thin; the boundary map includes all perceptual boundaries of the natural scene images, therefore achieving a high recall value. It appears evident that better boundary qualities will be found in the small scale signal. However, its small receptive fields cause extracting a significant level of noisy contours. So, precision, recall and thinning are provided by the small scale, and noise shortage is given by larger scales, mainly by the large scale. This analysis leads us to propose next assumptions to model the output stage of the PREEN architecture:

- Small scale signal provides all interesting boundaries for segmenting natural scenes.
- Medium and large scale signals enclose a reduced level of contour noise.
- An accurate boundary of a natural scene generates activity in all the scales.
- Output will be composed of thin and accurate boundaries, with a correct level of recall and generating activity in all the scales.

Basing upon these assumptions, PREEN model proposes a single boundary output signal resulting from the inter-scales competition for detecting boundaries with higher quality. Equation (17) shows the integration of the three scales through an inter-scale competition with major scale context dependent inhibition, ρ_{ijk} .

$$O_{ijk} = \left[\frac{U_{ijk}^{small} - \rho_{ijk} C_o \sum G_{pq}^{small} U_{pqk}^{small}}{A + U_{ijk}^{small} + \rho_{ijk} \sum G_{pq}^{small} U_{pqk}^{small}} \right]^+ \quad (17)$$

where U_{ijk}^{small} is the small scale CF hyper-complex activity, G_{pq}^{small} is a Gaussian kernel with a small scale deviation of σ_{small} , A is a decay constant, C_o is an inhibitory constant, and $[c]^+ = \max(0, c)$ and ρ_{ij} is the major scale context dependent inhibition gain, defined by equation (18).

$$\rho_{ijk} = \exp \left(-K_o \left(1 - \frac{1}{1 + \eta v_{ijk}} \right) \right) \quad (18)$$

$$v_{ijk} = \sum_{8\text{-neighbor}} U_{pqk}^{medium} + \sum_{8\text{-neighbor}} U_{pqk}^{large}$$

with K_o and η positive constants; v_{ijk} is the sum of the activities of scales medium and large in the 8-connected neighborhood of the consider position.

Equation (18) shows that ρ_{ijk} has a Gaussian profile, comprised between 0 and 1, with a shape dependent of K_o , and close to 0 when higher contour activity from larger scales. Equation (17) displays strong inhibition when the 8-connected neighborhood from larger scales has a high boundary activity and the boundary activity from the small scale will not

be affected if the neighborhood does not include contours. Image areas without contours in larger scales will sharply attenuate the boundary activity from the small scale.

3. Experimental results

We quantitatively compare our method against the Kokkinos et al.'s model (Kokkinos et al., 2008), that has been made available publicly. This model proposes a significant simplification of the BCS/FCS model proposed by Grossberg, Mingolla et al. (Grossberg et al., 1995) (Mingolla et al., 1999). The comparison to the Kokkinos et al.'s proposal (and to the BCS/FCS model by extension) is performed according to a common philosophy about propounding biologically inspired models. Differences among these models lie in the modeling of each stage and substantially in the color contribution of the PREEN model in the boundary detection processes in natural scenes. We will try to demonstrate that natural scenes have a chromatic component essential for defining and processing them. The comparison is accomplished with the boundary-based error, F-measure, which are the measure used in Kokkinos et al.'s paper. The measure computation has been made using the Matlab code supplied in next web page: <http://www.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/segbench/>. We have used all the 100 test images from the Berkeley Segmentation Dataset (Martin et al., 2001) and their human segmented images were taken as the ground truth to accomplish the F-measure values. Fig. 8-bottom shows a precision-recall curve (PR curve) with the average F-measure value of the boundaries detected in the all 100 test images. In Fig. 9 it can be observed processing examples for the PREEN architecture, including the F-measure values with their precision and recall as well as the position that would achieve in the ranking published in the Berkeley Segmentation Benchmark web page (BSDS, 2001).

Kokkinos et al. (Kokkinos et al., 2008) compared its model to the Canny edge detection algorithm using the images from the Berkeley dataset achieving better results in different scenarios. In their paper, they show the PR curves and the F-measure of the comparisons. Their better setting match with their BCS learned fine scale proposal, achieving an F-measure of 0.573 versus the 0.568 obtained by the Canny's algorithm. PREEN model achieves an F-measure with PR coordinates of 0.59 (0.68, 0.52), as it is shown in Fig. 8. So, it obtains better average result when detecting boundaries. Hence, PREEN boundary detection process achieves higher precision and recall values.

There are three notable differences between Kokkinos et al.' model and PREEN model. As previously said, the more significant one is the color contribution in the PREEN model. Another modeling difference is that PREEN model includes orientational competition, so PREEN obtains more accurate contours. A third significant difference is the output of the model. PREEN includes an inter-scale competitive stage as architecture output. This fusion provides us with an output enclosing all the better features of each scale. The substantial effect of color in the boundary detection process can be analyzed in the boundary detection algorithm ranking published in the Berkeley Segmentation Benchmark web page (<http://www.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/segbench/bench/html/algorithms.html>). The summary table shows different algorithms for grayscale and color processing. There are versions of algorithms for processing both grayscale and color images and it can be observed that color version obtains better results. Considering the average F-value achieved, PREEN architecture would gain the eighth position in the ranking

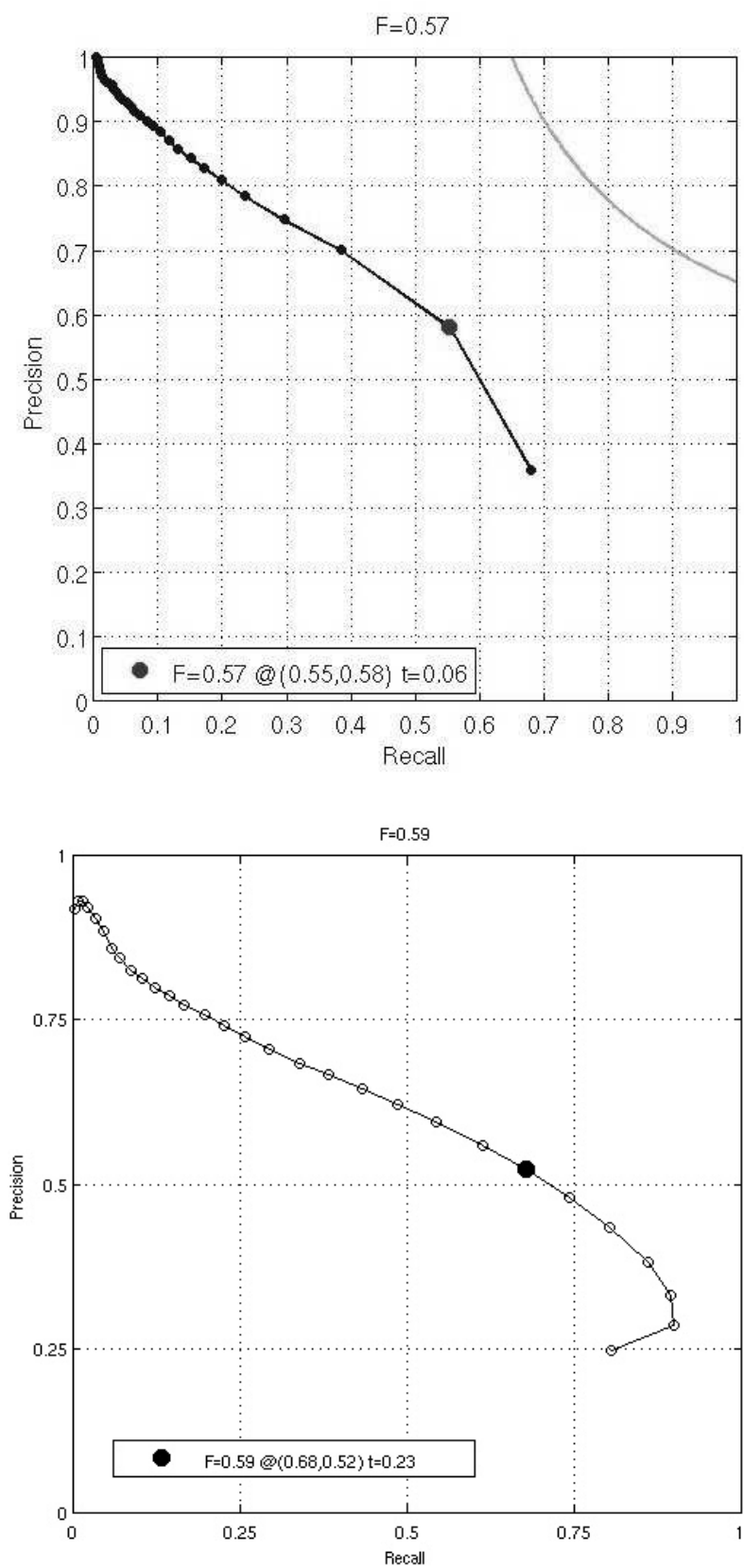


Fig. 8. Comparative results when processing all the 100 images from the test set of the Berkeley database. Top: Color gradient algorithm. Bottom: PREEN architecture.

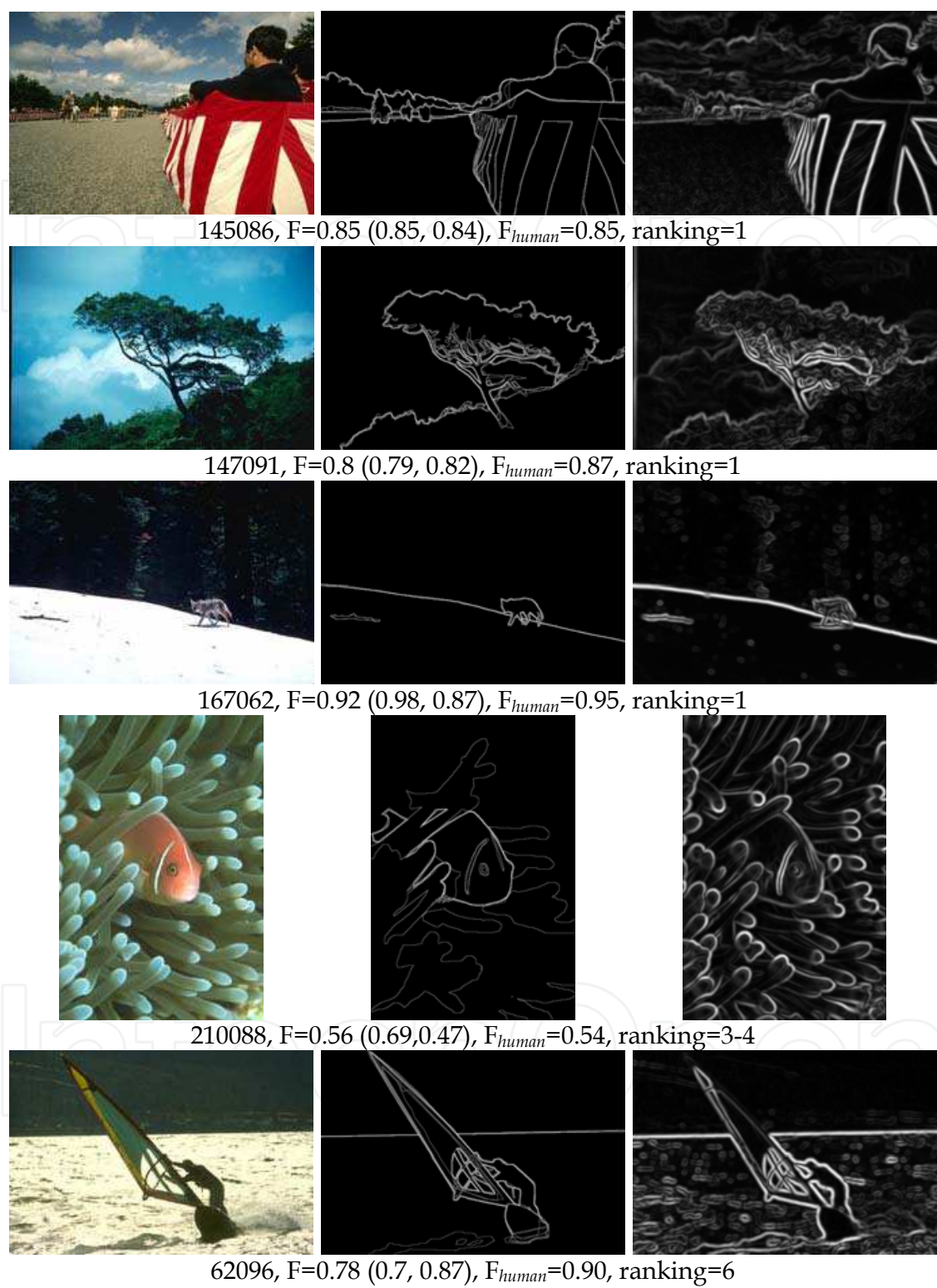


Fig. 9. Processing examples of PREEN architecture. Each example includes the image name in the Berkeley Segmentation Dataset, the F-measure value, the recall and precision coordinates, and the position that would achieve in the ranking published in the Berkeley Segmentation Benchmark web-site.

published in the BSDS web page, outperforming the Color Gradient algorithm. The PR curve of this method can be seen in Fig. 8-left (extracted from the aforementioned web page) comparing it to that of the PREEN model. Fig. 8 shows PREEN model achieves outstanding results in some benchmark images, overcoming in some cases the first algorithm in ranking. However, global result is not so excellent, locating the eighth position. This contradiction can be accurately analyzed considering the results shown in Fig. 4 c)-left. The surround effect determines the suppression of weak contours sited near other stronger. When processing image 86016 by the surround suppression PREEN model, we can observed lineal contours are drastically attenuated while round contour is reactivated. Apparently, this can be observed as a positive effect, since it affords the possibility to segment the scene into two regions with perceptual significance: the furrows area and the vegetation round area. However, segmentation performed by humans in the Berkeley database includes lineal contours as significant perceptual boundaries, which makes F-measure achieved by the PREEN architecture under these circumstances will lean to be lower, considerably diminishing the average F-measure value. Fig. 4 c)-right shows the opposite case, PREEN model causes the elimination of weak contours from inside tree branch, which makes boundaries to be finer considering the human segmentation.

4. Conclusion

In this paper we propose a model bio-inspired in the recurrent interactions of the early visual areas for detecting boundaries in color natural scenes, called PREEN model. To perform the comparative analysis, we have chosen BCS/FCS model (Grossberg et al., 1995) (Mingolla et al., 1999) and Kokkinos et al.'s model (Kokkinos et al., 2008) on the grounds of the similar aims. Images used are taken from the Berkeley Segmentation Dataset (BSDS, 2001). Important features of the proposed model are providing a simple and effective modeling, using even and odd oriented filtering, participation of surround suppression mechanisms in the complex cell stage, oriented and positional shunting competition, contour cooperative saliency extraction through bipole interactions, and a new output stage with inhibition among stages.

Test simulations over al the 100 test images from the Berkeley dataset and a comparative analysis have been included. Results achieved outperform those obtained by Kokkinos et al.'s model.

Furthermore, PREEN model has displayed outstanding results, overcoming the best algorithms published in the BSDS web page when processing some images of the database. In the analysis performed we have observed PREEN model favors boundaries with high perceptual significance attenuating the remainder boundaries. Thereby, PREEN model obtains low values in such images with human segmentations including boundaries of secondary importance, more germane to attentive mechanisms. We think if the database would have human segmentations with higher participation, the proposed model would achieve more favorable results. This higher participation, in our view, would mean remove or strongly attenuate secondary boundaries in human segmentations.

5. References

Antón-Rodríguez, M.; Díaz-Pernas, F. J.; Díez-Higuera, J. F.; Martínez-Zarzuela, M.; González-Ortega, D. & Boto-Giralda, D. (2009). Recognition of coloured and

- textured images through a multi-scale neural architecture with orientational filtering and chromatic diffusion. *Neurocomputing*, Vol. 72 (3713–3725). ISSN: 0925-2312.
- Berkeley segmentation dataset and benchmark (2001).
<http://www.cs.berkeley.edu/projects/vision/grouping/segbench/>.
 Last visited: September 2010.
- Daugman, J.G. (1980) Two-dimensional spectral analysis of cortical receptive field profiles. *Vision Research*, Vol. 20 (pp. 847–856). ISSN: 0042-6989.
- Grigorescu, C.; Petkov, N. & Westenberg, M.A. (2004). Contour and boundary detection improved by surround suppression of texture edges, *Image and Vision Computing*, Vol. 22, No. 8 (pp. 609–622). ISSN: 0262-8856.
- Grossberg, S. & Mingolla, E. (1985). Neural dynamics of form perception: Boundary completion, illusory figures and neon color spreading. *Psychological Review*, Vol. 92 (pp. 173–211). ISSN: 1939-1471.
- Grossberg, S., Mingolla, E. and Williamson, J. (1995). Synthetic aperture radar processing by a multiple scale neural system for boundary and surface representation. *Neural Networks*, Vol. 8 (pp. 1005–1028). ISSN: 0893-6080.
- Hodgkin, A.L. & Huxley, A.F. (1990). A quantitative description of membrane current and its application to conduction and excitation in nerve, *Bulletin of Mathematical Biology*, Vol. 52, No. 1-2 (pp. 25–71). ISSN: 0092-8240.
- Hubel, D. H. (1995). *Eye, Brain, and Vision*. Scientific American Library. W. H. Freeman, New York. ISBN: 978-0-7167-6009-2.
- Kokkinos, I.; Deriche, R.; Faugeras, O. & Maragos, P. (2008). Computational analysis and learning for a biologically motivated model of boundary detection. *Neurocomputing*, Vol. 71, No. 10-12 (Jun. 2008) (pp. 1798–1812). ISSN: 0925-2312.
- Kokkinos, I.; Deriche, R.; Papadopoulos, T.; Faugeras, O. & Maragos, P. (2007). Towards bridging the gap between biological and computational image segmentation. *INRIA Research Report: Thème BIO- Systèmes biologiques*, Vol. 6317. ISSN: 0249-6399.
- Landy, M.S. & Bergen, J.R. (1991). Texture segregation and orientation gradient. *Vision Research*, Vol. 31, No. 4 (pp. 679–693). ISSN: 0042-6989.
- Martin, D.; Fowlkes, C.; Tal, D. & Malik, J. (2001). A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics, *Proceedings of the 8th International Conference on Computer Vision*, vol. 2, pp. 416–423, ISBN: 0-7695-1143-0, Vancouver, Canada, July 2001.
- Mingolla, E.; Ross, W. & Grossberg, S. (1999). A neural network for enhancing boundaries and surfaces in synthetic aperture radar images. *Neural Networks*, Vol. 12 (499–511). ISSN: 0893-6080.
- Neumann, H. & Sepp, W. (1999). Recurrent V1–V2 interaction in early visual boundary processing, *Biological Cybernetics*, Vol. 91 (pp. 425–444). ISSN: 0340-1200.
- Neumann, H.; Yazdanbakhsh, A. & Mingolla, E. (2007). Seeing surfaces: The brain's vision of the world. *Physics of Life Reviews*, Vol. 4 (pp. 189–222). ISSN: 1571-0645.

- Ohta, Y.I.; Kanade, T. & Sakai, T. (1980). Color information for region segmentation. *Computer graphics and image processing*, Vol. 13 (222-241).
- Petkov, N. & Westenberg, M.A. (2003). Suppression of contour perception by band-limited noise and its relation to non-classical receptive field inhibition, *Biological Cybernetics*, Vol. 88 (pp. 236-246). ISSN: 1432-0770.
- Wilson, H. R.; Levi, D.; Maffei, L.; Rovamo, J. & De Valois, R. (1990). The perception of form: Retina to striate cortex. In: *Visual perception: The neurophysiological foundations*, L. Spillmann & J. S. Werner (Eds.), (pp. 231-272). San Diego: Academic Press, Inc. ISBN: 0-1265-7675-0.
- Zrenner, E.; Abramov, I.; Akita, M.; Cowey, A.; Livingstone, M. & Valberg, A. (1990). Color Perception: Retina to Cortex. In *Visual perception: The neurophysiological foundations*, L. Spillmann & J. S. Werner (Eds.), (Chap. 8). San Diego: Academic Press, ISBN: 0-1265-7675-0.

IntechOpen

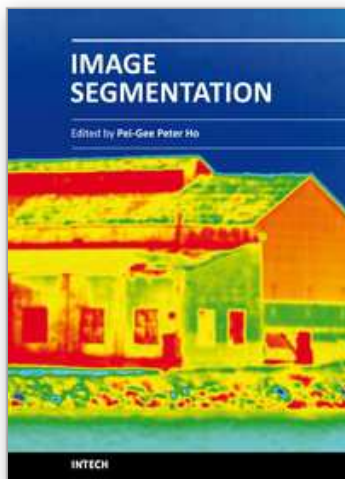


Image Segmentation

Edited by Dr. Pei-Gee Ho

ISBN 978-953-307-228-9

Hard cover, 538 pages

Publisher InTech

Published online 19, April, 2011

Published in print edition April, 2011

It was estimated that 80% of the information received by human is visual. Image processing is evolving fast and continually. During the past 10 years, there has been a significant research increase in image segmentation. To study a specific object in an image, its boundary can be highlighted by an image segmentation procedure. The objective of the image segmentation is to simplify the representation of pictures into meaningful information by partitioning into image regions. Image segmentation is a technique to locate certain objects or boundaries within an image. There are many algorithms and techniques have been developed to solve image segmentation problems, the research topics in this book such as level set, active contour, AR time series image modeling, Support Vector Machines, Pixon based image segmentations, region similarity metric based technique, statistical ANN and JSEG algorithm were written in details. This book brings together many different aspects of the current research on several fields associated to digital image segmentation. Four parts allowed gathering the 27 chapters around the following topics: Survey of Image Segmentation Algorithms, Image Segmentation methods, Image Segmentation Applications and Hardware Implementation. The readers will find the contents in this book enjoyable and get many helpful ideas and overviews on their own study.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Francisco J. Díaz-Pernas, Miriam Antón-Rodríguez, Isabel de la Torre-Díez, Mario Martínez-Zarzuela, David González-Ortega, Daniel Boto-Giralda and J. Fernando Díez-Higuera (2011). Surround Suppression and Recurrent Interactions V1-V2 for Natural Scene Boundary Detection, Image Segmentation, Dr. Pei-Gee Ho (Ed.), ISBN: 978-953-307-228-9, InTech, Available from: <http://www.intechopen.com/books/image-segmentation/surround-suppression-and-recurrent-interactions-v1-v2-for-natural-scene-boundary-detection>

INTECH
open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

www.intechopen.com

IntechOpen

IntechOpen

© 2011 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](https://creativecommons.org/licenses/by-nc-sa/3.0/), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen