

1991-02

Preattentive Texture Segmentation and Grouping by the Boundary Contour System

<https://hdl.handle.net/2144/2063>

Downloaded from DSpace Repository, DSpace Institution's institutional repository

**PREATTENTIVE TEXTURE SEGMENTATION
AND GROUPING BY THE
BOUNDARY CONTOUR SYSTEM**

Dan Cruthirds, Alan Gove, Stephen Grossberg, and Ennio Mingolla

February, 1991

Technical Report CAS/CNS-TR-91-008

Permission to copy without fee all or part of this material is granted provided that 1. the copies are not made or distributed for direct commercial advantage, 2. the report title, author, document number, and release date appear, and 3. notice is given that copying is by permission of the BOSTON UNIVERSITY CENTER FOR ADAPTIVE SYSTEMS AND DEPARTMENT OF COGNITIVE AND NEURAL SYSTEMS. To copy otherwise, or to republish, requires a fee and/or special permission.

Copyright © 1991

Boston University Center for Adaptive Systems and
Department of Cognitive and Neural Systems
111 Cummington Street
Boston, MA 02215

PREATTENTIVE TEXTURE SEGMENTATION AND GROUPING BY THE BOUNDARY CONTOUR SYSTEM

Dan Cruthirds, Alan Gove, Stephen Grossberg & Ennio Mingolla
Center for Adaptive Systems and Graduate Program in Cognitive and Neural Systems
Boston University, 111 Cummington Street, Boston, MA 02215.

ABSTRACT: An improved Boundary Contour System (BCS) neural network model of preattentive vision is applied to two images that produce strong "pop-out" of emergent groupings in humans. In humans these images generate groupings collinear with or perpendicular to image contrasts. Analogous groupings occur in computer simulations of the model. Long-range cooperative and short-range competitive processes of the BCS dynamically form the stable groupings of texture regions in response to the images.

1. Introduction: Segmenting image regions even in the absence of luminance differences has been a persistent problem in vision research. The present article describes an improved BCS model that is able to segment multi-element textures in a manner analogous to texture segmentation by human observers. Digital grey-scale CCD images of two Beck-style displays (Beck, Prazdny, and Rosenfeld, 1983) are used as inputs to the system.

2. Description of the Model: The neural network model used is a refinement of the Boundary Contour System developed by Grossberg and Mingolla (1985a, 1985b, 1987) through an analysis of biological vision. Along with its companion Feature Contour System (FCS), the BCS has been used to explain a variety of visual effects, including brightness perception, boundary completion, hyperacuity, depth perception, neon color spreading, and binocular rivalry (Grossberg, 1987; Grossberg and Mingolla, 1985a, 1985b; Grossberg and Todorovic, 1988). The BCS model locates and completes boundaries that delimit regions for filling-in with featural (color and brightness) signals. The model used for this work is similar to the BCS described in Grossberg and Mingolla (1987). The improved model adds an off-channel in the early processing, a new version of the bipole filters, and another stage in the CC Loop (defined below). The role of each processing stage is indicated in Figure 1 and explained as follows:

Stage 1 (ON Cells and OFF Cells): This stage is accomplished by two shunting center-surround systems. The first, an on-center off-surround network, corresponds to an "ON" channel of the visual pathway. Likewise, the second shunting network, with an off-center and on-surround, corresponds to an "OFF" channel. In each case the equilibrium state of the dynamical system contains both a DOG (Difference of Gaussians) term, which detects contrast differences, and a term which compensates for the level of illumination, thereby discounting the illuminant. The two networks differ in sign in their response to a given light-to-dark (left-to-right) step transition, as the ON channel responds positively on the left side of the step, and the OFF channel responds positively on the right side of the step (negative outputs are set to zero). Both channels are tuned to give a null response to uniformly illuminated areas.

Stage 2 (Simple Cells): The oriented simple cells use both the ON and OFF channels to gauge oriented contrast differences at each image location. An edge elicits a strong response in the ON channel to one side and a strong OFF channel response to the other side. Oriented outputs are calculated for twelve orientations across 180 degrees. The resulting spatial representation contains all twelve oriented outputs for every point in the original image.

Stage 3 (Complex Cells): The next level of complex cells compensates for direction of contrast by combining the rectified outputs of dark-to-light and light-to-dark simple cells at each orientation. A shunting filter across space and orientation sharpens the response, thereby creating thinner boundaries.

Stage 4 (Hypercomplex cells: First Competitive Stage): The first stage of the cooperative-competitive feedback net, or CC Loop, consists of a competition within orientation and across spatial position. This plays the role of an endstopping operation that converts complex cells into hypercomplex cells. Hypercomplex cells receive inputs from the oriented complex cells as well as positive feedback signals from long-range cooperative processes (described below). All cells also receive a tonic input which energizes disinhibitory activations of cells whose competitors are inhibited by endstopped signals.

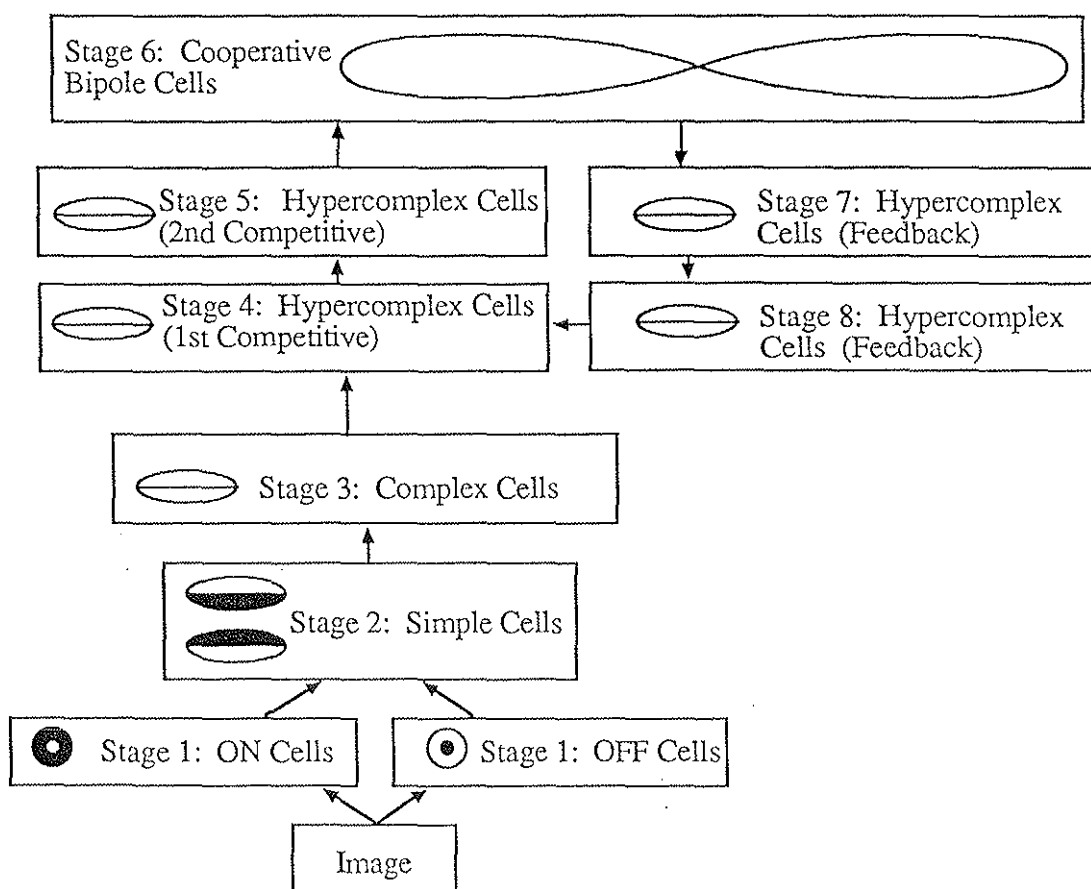


Figure 1: A schematic diagram of the stages of the revised BCS model.

Stage 5 (Hypercomplex cells: Second Competitive Stage): This stage is complementary to the prior stage, in that it computes a competition within position but across orientation. Here, perpendiculars to orientations inhibited in the prior stage are disinhibited. Thus, along the sides and at the ends of lines, signals perpendicular to those received from the complex cells (Stage 3) flank the bottom-up orientation signals. These “end cuts” aid in perceptual groupings involving line ends. This and the previous stage define the competitive portion of the CC Loop.

Stage 6 (Cooperative Bipole Cells): The cooperative portion of the CC Loop is performed at this stage by bipole cells that act like long-range statistical AND gates. Cells with the predicted bipole properties have been described by von der Heydt, Peterhans, and Baumgartner (1984) and Peterhans and von der Heydt (1989). In order for a horizontally oriented cooperative bipole cell to fire, both the left and right receptive fields of the cell need to receive input signals from the hypercomplex cells of Stage 5. When a bipole cell fires, it sends a top-down signal through Stages 7 and 8 to the hypercomplex cells of Stage 4, where it is combined with bottom-up information. This type of boundary completion can occur simultaneously across all orientations at all positions.

Stages 7 and 8 (Hypercomplex Cells): Before cooperative signals are sent to the first competitive stage, a competition homologous to the first and second competitive stages takes place in order to pool and sharpen the signals that are fed back.

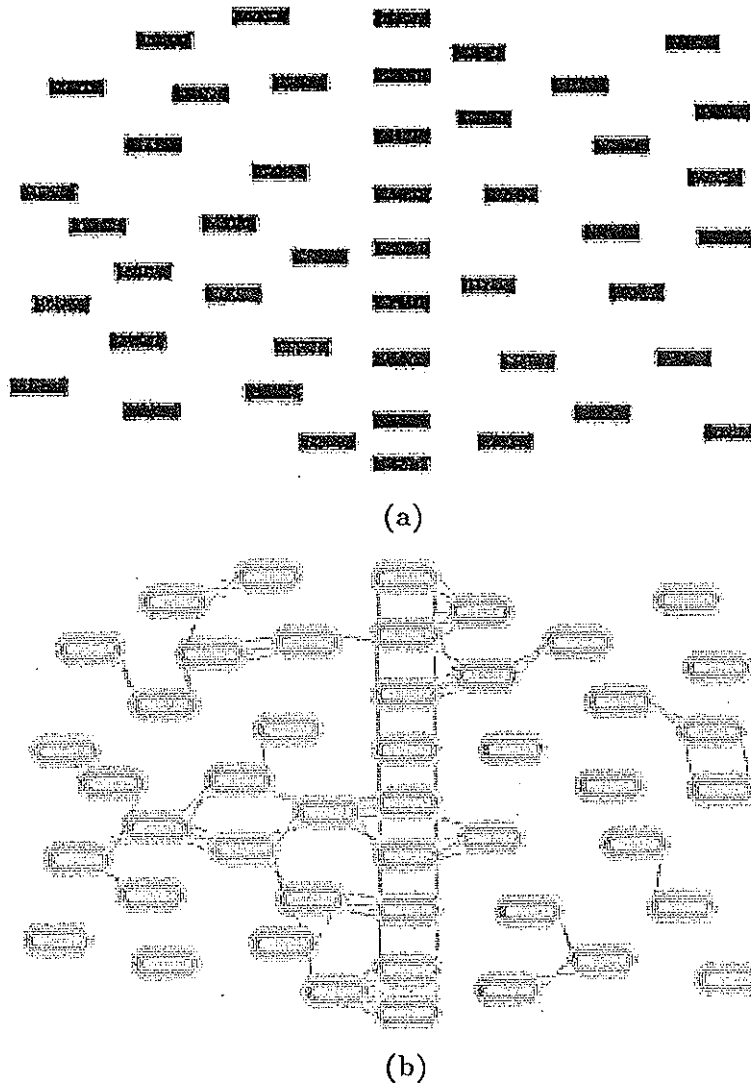
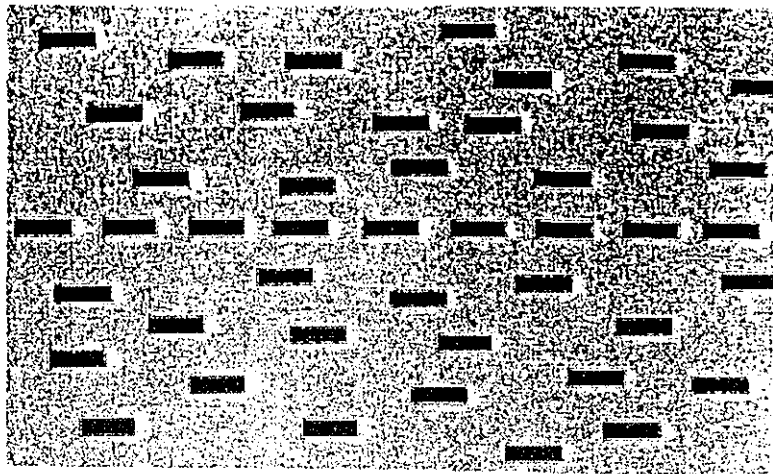
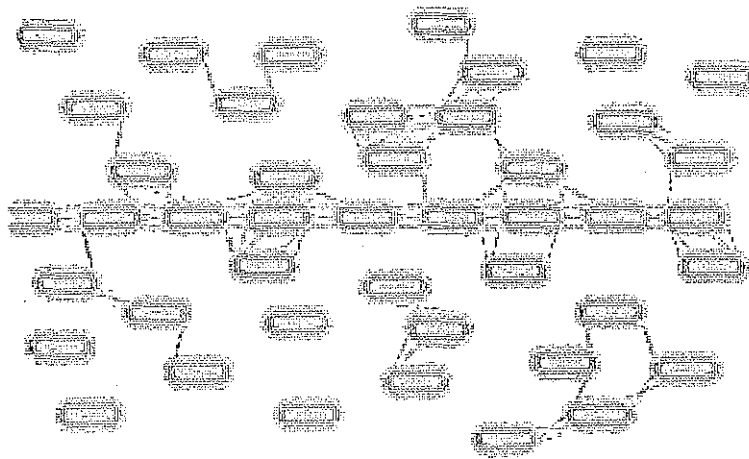


Figure 2 (a) At appropriate viewing distances the vertical row formed by the aligned bars in the center of the figure preattentively “pops out” for human observers. (b) The output of the refined BCS builds links between those bars whose juxtaposition is conducive to long-range grouping. While these links do not correspond to visible contrast in the image, they mimic the powerful groupings that lead to pop-out in (a).



(a)



(b)

Figure 3 (a) As in Figure 2, aligned bars form a preattentively segmented long horizontal structure. (b) The model detects and enhances the horizontal “stripe” formed by the aligned bars in the middle. Note that the cooperative groupings are both “long-range” often spanning distances that are several times the diameters of individual elements, and “sharp” in that they do not get fuzzier as grouping distances get longer.

3. Results and Discussion: Figures 2a and 3a display scenes of rectangular elements (bars) that are pseudorandomly distributed, except for collinear arrangement of bars that run vertically (2a) or horizontally (3a) through the middle of the images. Figures 2b and 3b show the outputs of the hypercomplex cells of the second competitive stage (Stage 5) at equilibrium. Figure 4 displays enlarged views of a small region of Figure 2b, in order to clarify the distributions of signals from hypercomplex cells of various orientations. Note that the fuzzy structures surrounding the bar region are results of the “endcut” process. The structures connecting the representation of certain of the neighboring bars are *emergent boundaries* that indicate perceptual groupings of bars into higher-order structures. Sufficiently strong groupings, such as those running horizontally or vertically throughout the display, can be useful both for segmentation and as inputs to a process that gener-

ates “pop-out” or separation of figure from ground. Note that many of the groupings of Figures 2b and 3b occur through regions which yield no sense of enhanced lightness or darkness when one views the corresponding regions of Figures 2a or 3a. Boundaries that are the outputs of the BCS are said to be “invisible” for this reason. The BCS boundaries that surround each individual bar *do succeed* in trapping different values of featural signals on the interior of each bar boundary than on the exterior, thereby rendering *those* boundaries visible in the sense of producing a consciousness of a difference in light and dark at the two sides of the boundary. The theory on which the present work is based has clarified why invisible boundaries – so often overlooked in other approaches precisely because they *are* invisible – can powerfully influence which objects in a scene are rapidly segmented from their background and recognized (Grossberg and Mingolla, 1985a, 1985b; Grossberg, 1987).

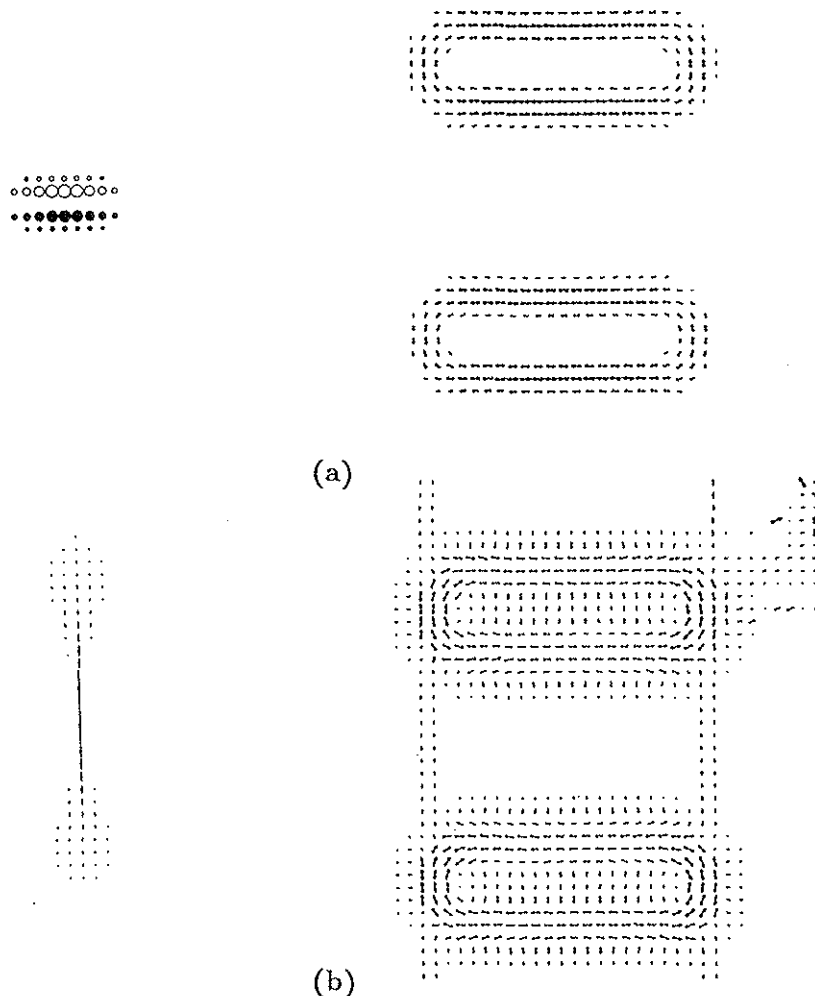


Figure 4 (a) Simple cell kernel and the output of such kernels at all orientations before feedback is initiated for an enlarged portion of Figure 2a. (d) Bipole cell kernel and the output of Stage 5 at equilibrium.

4. **Conclusions:** The groupings of the model shown in this paper attempt to emulate one of the most difficult and fundamental of human visual competencies: the dynamic, automatic synthesis of the perceptually salient *units* in a visual scene. Because those units are so often topologically unconnected in the luminance domain, mammalian visual systems have evolved strategies for rapid, preattentive groupings of a scene using networks that form locally connected, massively parallel

topographic maps. In attempting to emulate this competence in a computer-implemented model, we have identified and simulated a number of local circuits for keeping a large positive feedback process in check. As important as the groupings that do survive is the active suppression of those potential groupings that do not. While the possibility of grouping at any orientation over a variety of distances forces a high degree of uncertainty or fuzziness at several model stages, the displayed output contains emergent boundaries that are hyperacute, in the sense that they are thinner, as in Figure 4, than the input bars in the luminance domain of Figures 2a and 3a that induce them, and the simple cell receptive fields that detect these inputs. Moreover the thickness of the emergent boundaries is relatively independent of their length. Like their perceptual analogs, the emergent groupings are not blurred, as they would be if they were generated by a low spatial frequency filter. The present investigations clarify how long range perceptual effects can be generated by local operations, and illustrate how automatic processes can achieve competencies that many previous approaches have attributed to top-down or "cognitive" factors.

Acknowledgements: This research was supported in part by the Air Force Office of Scientific Research (AFOSR 90-0175) and Hughes Aircraft Company (SK-902369-SDB).

References

- Beck, J., Prazdny, K. and Rosenfeld, A. (1983). A theory of textural segmentation. In. J. Beck, B. Hope, and A. Rosenfeld (Eds.), *Human and Machine Vision*. New York: Academic Press.
- Grossberg, S. (1987). Cortical dynamics of three-dimensional form, color, and brightness perception, II: Binocular theory, *Perception and Psychophysics*, **41**, 117-158.
- Grossberg, S. and Mingolla, E., (1985a). Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading, *Psychological Review*, **92**, 173-211.
- Grossberg, S. and Mingolla, E., (1985b). Neural dynamics of perceptual grouping: Textures, boundaries, and emergent segmentations, *Perception and Psychophysics*, **38**, 141-171.
- Grossberg, S. and Mingolla, E., (1987). Neural dynamics of surface perception: Boundary webs, illuminants, and shape-from-shading, *Computer Vision, Graphics, and Image Processing*, 1987, **37**, 116-165.
- Grossberg, S. and Todorović, D., (1988). Neural dynamics of 1-D and 2-D brightness perception: A unified model of classical and recent phenomena, *Perception and Psychophysics*. **43**, 241-277.
- Peterhans, E. and von der Heydt, R. (1989). Mechanisms of contour perception in monkey visual cortex. II. Contours bridging gaps. *The Journal of Neuroscience*, **9**, 1749-1763.
- von der Heydt, R., Peterhaus, E., and Baumgartner, G., (1984). Illusory contours and cortical neuron responses, *Science*, **224**, 1260-1262.