Proto-object based contour detection and figure-ground segmentation

Brian Hu, Rüdiger von der Heydt, and Ernst Niebur

Summary. Segmenting an image into regions corresponding to objects requires contour detection and figure-ground segmentation. Border-ownership selective cells [1] are well-suited for this task, as they encode where an object is located relative to an edge in their receptive fields. We develop a recurrent neural model using border-ownership selective (B) cells and grouping (G) cells whose activity represents proto-objects based on the integration of local feature information. G cells send feedback connections to the B cells that caused their activation, which when combined with local inhibitory competition, helps to determine figure-ground assignment. We evaluate our model on the Berkeley Segmentation Dataset (BSDS-300) [2], and achieve performance comparable to recent state-of-the-art computer vision approaches. These approaches are computationally expensive due to their requirement for training and typically provide only limited insight into the underlying functional mechanisms. In comparison, our model (1) is biologically plausible, (2) is fully image computable, and (3) does not require any training.

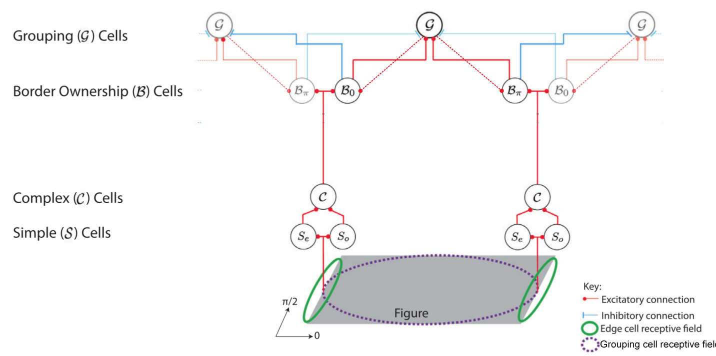

Figure 1: Grouping model

Methods. We extend a purely feedforward model of proto-object based saliency [3] to include recurrent connections, implemented by an iterative update rule. Previous models of perceptual organization [4] included recurrent connections but could not be applied to natural images. We used the original BSDS grayscale images and then successively downsampled in half-octaves to create an image pyramid with ten levels. At each level, we apply an edge detector based on a model of simple cells with push-pull inhibition [5]. The same set of feedforward and feedback grouping operations is then applied to the edge input independently at each level of the pyramid to achieve scale invariance. G cells receive input from B cells and, inspired by Gestalt principles such as closure, continuity, and convexity, their receptive fields integrate global context information about proto-objects in the scene by summing cocircular edge signals. Figure 1 shows the structure of our model. Interactions between G cells of different scales operates in a coarse-to-fine manner and allows global context to influence figure-ground assignment. B cells are driven by orientation selective (complex) cells and receive modulatory feedback from G cells which makes them border-ownership selective. Additionally, B cells with opposite side-of-figure preferences compete via local inhibition (not shown in figure). Contour detection and figure-ground assignment results are computed from the population of B cells at the highest resolution level. B cell activity is converted into a population vector code, where the magnitude of the resulting vector represents edge strength, and the direction of the vector provides a continuous figure-ground orientation label. For model evaluation, we apply publicly available benchmarking code [2, 6] to the standard set of 100 test images from the BSDS dataset and report F-scores for contour detection and mean accuracy for figure-ground assignment.

Results. Figure 2 shows contour detection and figure-ground assignment results for a sample image. Our model can accurately detect contours and assign figure-ground relationships to these contours based on the objects present in the image. Evaluation across the full set of BSDS test images resulted in an F-score of 0.60 on the contour detection benchmark and a mean accuracy of 69.7% on the figure-ground assignment benchmark. These results are comparable to the best results obtained so far on contour

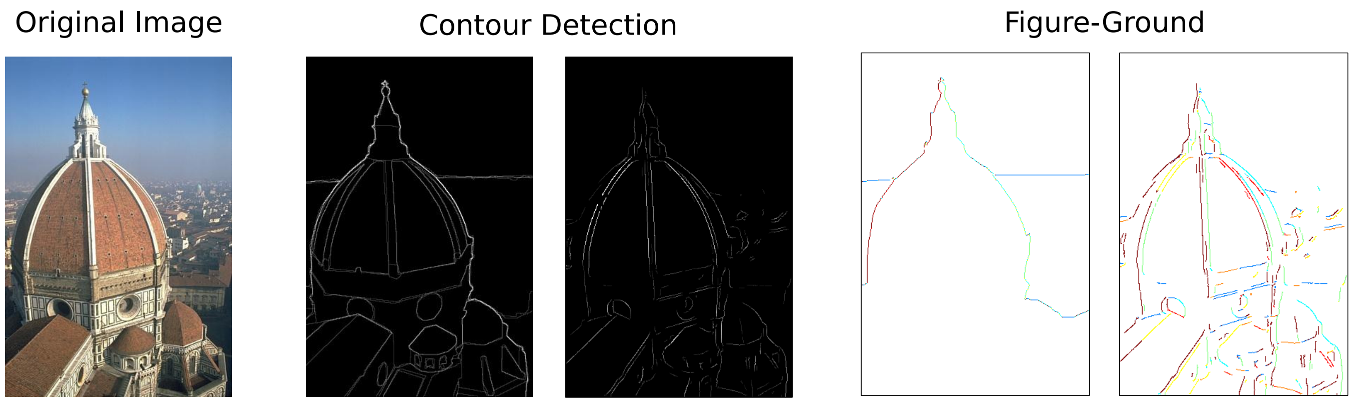

Figure 2: Results on an example image from the BSDS dataset. A) The original color image. B) The human ground truth (left) and model results (right) for contour detection. C) The human ground truth (left) and model results (right) for figure-ground assignment. Each color represents one of eight possible binned figure-ground orientations.

detection [7, F-score 0.68] and figure-ground assignment [6, mean accuracy 74.7%]. Unlike approaches based on machine learning, our model is biologically plausible with computational mechanisms that have direct neural correlates. Future work will compare our model results with recent experimental results in which border ownership coding was studied in natural scenes [8].

[1] H. Zhou, H. S. Friedman, and R. von der Heydt. *J. Neurosci.*, 20(17):6594–6611, 2000. PMID: 10964965.

[2] D. Martin, C. Fowlkes, D. Tal, and J. Malik. A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics. In *Proc. 8th Int'l Conf. Computer Vision*, volume 2, pages 416–423, July 2001.

[3] A. F. Russell, S Mihalas, R. von der Heydt, E. Niebur, and R. Etienne-Cummings. *Vision Research*, 94:1–15, 2014.

[4] E. Craft, H. Schütze, E. Niebur, and R. von der Heydt. *Journal of Neurophysiology*, 97(6):4310–26, 2007. PMID: 17442769.

[5] George Azzopardi, Antonio Rodríguez-Sánchez, Justus Piater, and Nicolai Petkov. *PloS one*, 9(7):e98424, 2014.

[6] Ching Teo, Cornelia Fermuller, and Yiannis Aloimonos. Fast 2d border ownership assignment. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5117–5125, 2015.

[7] Pablo Arbelaez, Michael Maire, Charless Fowlkes, and Jitendra Malik. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(5):898–916, May 2011.

[8] Jonathan R Williford and Rudiger von der Heydt. *Journal of Vision*, 14(10):588–588, 2014.