



Consistent and robust determination of border ownership based on asymmetric surrounding contrast

Ko Sakai^{a,*}, Haruka Nishimura^b, Ryohei Shimizu^a, Keiichi Kondo^a

^a Department of Computer Science, University of Tsukuba, Tsukuba, Japan¹

^b Institute of Intelligent Systems Research, National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan

ARTICLE INFO

Article history:

Received 18 July 2010

Received in revised form 19 April 2012

Accepted 18 May 2012

Keywords:

Computational modeling

Middle vision

Perceptual organization

Receptive field

Shape and contour

ABSTRACT

Determination of the figure region in an image is a fundamental step toward surface construction, shape coding, and object representation. Localized, asymmetric surround modulation, reported neurophysiologically in early-to-intermediate-level visual areas, has been proposed as a mechanism for figure–ground segregation. We investigated, computationally, whether such surround modulation is capable of yielding consistent and robust determination of figure side for various stimuli. Our surround modulation model showed a surprisingly high consistency among pseudorandom block stimuli, with greater consistency for stimuli that yielded higher accuracy of, and shorter reaction times in, human perception. Our analyses revealed that the localized, asymmetric organization of surrounds is crucial in the detection of the contrast imbalance that leads to the determination of the direction of figure with respect to the border. The model also exhibited robustness for gray-scaled natural images, with a mean correct rate of 67%, which was similar to that of figure-side determination in human perception through a small window and of machine-vision algorithms based on local processing. These results suggest a crucial role of surround modulation in the local processing of figure–ground segregation.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

The surround modulation reported in early-to-intermediate-level visual areas is generally thought to play a crucial role in figure–ground segregation (Fitzpatrick, 2000; Lamme, 1995; Super, Spekreijse, & Lamme, 2001; Super, van der Togt, Spekreijse, & Lamme, 2003; Xiao, Raiguel, Marcar, Koenderink, & Orban, 1995; Zipser, Lamme, & Schiller, 1996) and in the determination of border ownership (BO) (Nishimura & Sakai, 2004; Sakai & Nishimura, 2006) that defines the direction of a figure along the border. Asymmetry of the surround structure is the essence of the mechanism. A facilitatory region on one side of the classical receptive field (CRF), with a suppressive region on the other side, is a simple example of a surround structure enabling figure–ground segregation. In this case, the neuronal firing rate is increased when a figure is projected onto the facilitatory side, and decreased when it is projected onto the alternate side, generating BO selectivity, i.e., BO can be computed as long as the surround modulation is imbalanced with respect to the CRF.

The crucial question concerning this type of surround modulation model is whether such unstructured surround regions

result in a consistent BO with respect to a variety of stimuli with distinct shapes. A BO-selective cell should exhibit a stronger response when a figure is presented on its preferred side versus the alternative side, regardless of stimulus shape, i.e., a cell coding BO should be shape invariant. For example, a cell would not be considered selective for BO if it responded stronger when a cup is presented on the right side of its CRF rather than on the left side, but weaker when a pen is presented on the right side versus the left side. Although BO-selective cells, identified *in vivo*, have only been tested using three simple types of stimulus (a single square, a C-shaped figure, and two overlapping squares), it has been assumed that many of these cells would be shape invariant. Whether these cells are indeed selective for BO needs to be judged by testing consistency among a wide variety of stimuli. Though such shape invariance is difficult to test physiologically, it is possible to use computational modeling and simulations to evaluate how models of surround modulation can yield consistent and robust BO determination.

To address the question of consistency, we developed pseudorandom block stimuli that approximate all possible shapes under certain constraints, and analyzed the consistency of a surround modulation model. The present model used randomly generated surround regions that approximate all possible regions in terms of their location, size, and shape under certain constraints, in order to enable a thorough analysis of the consistency. Note that the surrounds are not limited to being asymmetric, and they are not

* Corresponding author.

E-mail address: sakai@cs.tsukuba.ac.jp (K. Sakai).

¹ <http://www.cvs.cs.tsukuba.ac.jp>.

designed to achieve BO selectivity or to replicate the physiological experiments (Qiu & von der Heydt, 2007; Zhou, Friedman, & von der Heydt, 2000). It is natural to expect that such randomization may not evoke high consistency. We performed simulations of the model using the pseudorandom block stimuli, to evaluate the consistency of the model. We suspected that a model cell with a simple mechanism of surround modulation would exhibit consistency only for a limited number of stimuli whose shapes fall into its particular set of surround regions, and would exhibit inconsistency for other stimuli. Surprisingly, our simulation results indicated high consistency for the pseudorandom block stimuli, suggesting the robustness of the surround modulation mechanism with respect to stimulus shape. To identify the characteristics of the surround regions that are crucial for consistent shape-invariant BO determination, we performed a computational analysis of the surround structures. Physiological studies have indicated that, in most cases, surround regions are localized asymmetrically with respect to their CRF centers, and that suppression is dominant over facilitation (Jones, Wang, & Sillito, 2002; Walker, Ohzawa, & Freeman, 1999). We expected that such weak constraints on the surround modulation would result in the robust determination of BO. We applied a reverse-correlation technique to the model cells in order to extract the characteristics of the surround regions that are necessary for the consistent determination of BO. The model was constructed using localized surrounding regions generated randomly. Therefore, unconditional averaging of the regions resulted in a uniform distribution over the surrounding space. The conditional averaging of the regions of model cells with a specific consistency reveals the characteristics of the surround structure that are essential for the consistency. We used reverse correlation on groups of model cells with different degrees of consistency, to identify those characteristics of surround regions that are necessary for highly consistent BO selectivity. Note that the randomness of surround regions in the present model enabled the application of the reverse-correlation technique. Because the previous models (e.g. Sakai & Nishimura, 2006) assigned facilitatory and suppressive regions to the preferred and nonpreferred sides, respectively, unconditional averaging results in a nonuniform distribution (i.e., facilitation on the preferred side). This nonuniformity violates the prerequisite of the reverse-correlation analysis, and thus this technique is not applicable to the previous models. The reverse-correlation analysis of the present model showed that the surrounds with high consistency showed facilitation on the preferred side and suppression on the alternate side. This result suggests that the detection of contrast imbalance between the sides is the crucial aspect of surround modulation which enables high consistency in BO determination. Further quantitative analyses revealed that surrounds with higher consistency have stronger suppression, with more complementary localization of suppression and facilitation with respect to the border. These characteristics (the strong suppression and the localization) are consistent with those of the surround modulation reported in physiological experiments (Jones, Grieve, Wang, & Sillito, 2001; Walker et al., 1999).

The second question concerns the behavior of a population of model cells in the consistent determination of BO. Specifically, we wondered about the number of model cells that determine the correct (consistent) BO for each stimulus and about the nature of the stimulus characteristics that tend to evoke correct responses in a great number of model cells. A previous neurophysiological study showed that each BO-selective cell exhibits some degree of incomplete cue invariance, i.e., a cell shows a clear BO for some stimuli, but yields an ambiguous result for others (Zhou et al., 2000). It is crucial to assess the incomplete cue invariance of the model. To examine the responses of a set of model cells to each stimulus (for the evaluation of the behavior of the model as a group), we performed population analyses of the model cells.

Specifically, we examined the manner in which the population activity changed depending on the stimulus, and whether a set of model cells, as a whole, was capable of determining consistent BO for the block stimuli. Furthermore, we identified the stimulus characteristics that tended to evoke consistent determination of BO. Our analysis revealed that the numbers of blocks and edges on the nonpreferred side have negative effects, whereas those on the preferred side have positive effects in terms of the consistency (easiness) of stimulus. This suggests that, for such complementary organization, suppression on the nonpreferred side and facilitation on the preferred side is important for the surround structure in the consistent determination of BO.

The third question was whether the consistency of the model corresponds to human perception. The determination of the rank of stimuli regarding its consistency in BO determination enables the comparison between model performance and human perception. It is expected that stimuli that achieve higher consistency for the model will result in higher accuracy and shorter reaction times for human subjects judging BO. We therefore performed psychophysical experiments and confirmed the presence of such a relation between consistency and human perception.

Finally, the robustness of the model was examined using natural images. Tests performed using natural images clarify whether the surround contrasts that exist in these images are sufficient to evoke BO that is consistent with human perception. The model based on surround modulation computes BO from surround contrasts that are not limited to those evoked by contours, but include those evoked by the surround context, such as luminance, texture, and shading. It is crucial to investigate whether the wide variety of surround contrasts present in natural images evokes BO. We tested the model using a set of natural images from the Berkeley Segmentation Dataset (Martin, Fowlkes, Tal, & Malik, 2001). Specifically, we computed the BO of the images from the dataset, and compared the results with the human-marked BO of the dataset. Our model was in good agreement with human perception, indicating that the surround contrasts that exist in natural images are sufficient to evoke BO that is consistent with human perception. Natural images are extremely challenging for the contour-based analysis of BO, as the extraction of continuous, reliable contours from these images is almost impossible. In contrast, our model based on the balance of surrounding contrast is not limited by this problem.

2. Methods

2.1. Model architecture

To signal the BO direction, models that are based on surround suppression and facilitation modulate the response from the CRF according to their surround context. The model used in the present study shares this fundamental idea, with major extensions from the previous models, such as realistic and diverse surround modulation, to ensure physiological plausibility and computational thoroughness, with an intention of processing arbitrary figural shapes.

The model consisted of three stages: local contrast detection using Gabor filters, surround modulation with iso-orientation suppression and cross-orientation facilitation, and determination of BO based on the modulation of the CRF response by its surround contrast. Because the focus of interest was the spatial organization of the surround modulation and its diversity, we used a filter-based model rather than a biophysically detailed model with realistic temporal characteristics. An illustration of the model flow is shown in Fig. 1. For the simulations using block stimuli, we limited ourselves to determining BO exclusively in the horizontal directions, for the sake of simplicity. For the simulations using natural images, we extended the model to

Download English Version:

<https://daneshyari.com/en/article/405503>

Download Persian Version:

<https://daneshyari.com/article/405503>

[Daneshyari.com](https://daneshyari.com)