NEUROPHYSIOLOGICAL EVIDENCE FOR IMAGE SEGMENTATION AND MEDIAL AXIS COMPUTATION IN PRIMATE V1

Tai Sing Lee
tai@arcadia.harvard.edu

Department of Brain and Cognitive Sciences
MIT, Cambridge, MA 02139, U.S.A.

Division of Applied Sciences
Harvard University, Cambridge, MA 02138, U.S.A.

ABSTRACT
We studied the spatio-temporal behaviors of V1 neurons in awake rhesus monkeys in response to a set of static texture images. We found that V1 neurons are sensitive to luminance contrast, texture contrast and the global context of an image. The spatio-temporal response profiles of these neurons exhibit a gradual spatial contraction of the boundary signals and a simultaneous smoothing of the surface signals within a texture-defined figure. Furthermore, a sharp response peak is often observed when a cell’s classical receptive field is located at the center of the figures. This evidence suggests that V1 may play a role in the computation of the surface boundary and the medial axis of objects present in the visual scene.

INTRODUCTION
Recent computational models for image segmentation emphasize the coupling of two concurrent processes: surface interpolation and boundary detection[1,2].
These two dynamic processes, characterized by the gradual contraction of the boundary signals and the simultaneous smoothing of the surface signals, can be understood in a Bayesian framework for surface inference [3]. We and others [3,4] have postulated that such processes could be implemented by the neural machinery of the primary visual cortex, area V1. In this experiment, we studied the dynamics of V1 neurons in response to a set of static texture images with a view to understanding the cortical processes underlying image segmentation.

**METHODS**

The following three types of stimuli were presented to two awake behaving rhesus monkeys: (a) a boundary defined by two regions of contrasting texture, (b) a texture strip subtending 4° of visual angle on a background of contrasting texture, and (c) a 4° × 4° texture square on a background of contrasting texture. The texture within each region was composed of small randomly positioned lines of uniform orientation. The texture of the contrasting region was composed of lines of the orthogonal orientation. Squares and strips of various widths were also tested. Examples of the stimuli used are shown in Figure 1:

![Type a: Boundary](image) ![Type b: Strip](image) ![Type c: Square](image)

Figure 1: The “positive” stimuli for a vertically oriented cell are shown here. Both “positive” and “negative” stimuli were used in the experiment: a “positive” stimulus is one in which the figure is composed of the texture of the preferred orientation and a “negative” stimulus is one in which the figure’s texture is of the orthogonal orientation.

In each trial, a full screen sample stimulus was presented for 330 msec while the monkey fixated within a 0.3° × 0.3° fixation window. Each stimulus was presented at a randomized series of sampling positions relative to the cells’ classical receptive fields (CRF). The data samples were taken at 0.5° steps over a 12° range, in a direction perpendicular to the texture boundary across the middle of the texture-defined figure. The boundary was parallel to the preferred orientation of the cells.

The receptive fields of the cells described here were located between 3-6 degree in eccentricity and had diameter less than 0.5° visual angle.
**Results**

We recorded from 214 neurons in two rhesus monkeys. Of these, 120 units exhibited sharp and sustained responses to texture boundaries. Fifty-eight of these units were tested with texture squares, and 78 with texture strips.

The neuronal responses for the first 40-65 msec are characterized by responses to local features. The cells respond briskly and transiently within the whole texture region. By 120 msec after stimulus onset, the responses within the texture region decay to a uniform level whereas the responses at the texture boundary are relatively sustained, resulting in a gradual sharpening of the boundary signals as shown in Figure 2.

![Figure 2](image-url)

**Figure 2:** A cell’s response gradually contracts toward the boundary of a “type a” stimulus, exhibiting the gradual sharpening effect. The abscissa is the distance in visual angle from the RF to the texture boundary. At 0°, RF overlaps with the texture boundary. The texture in the region to the right of the boundary is of the preferred orientation of the cell. The solid line indicates the mean firing rate within a 15 msec window. The dash lines provide the standard error envelope. The dots on the solid line are the data points.

V1 neurons also developed similar sharp boundary signals in response to the squares and the strips. But the response within the square is significantly more sustained, showing the figural enhancement effect [5]. In addition, a sharp response peak was often observed when the cells’ CRF overlapped the center of the strip or the square. It emerged simultaneously with the emergence and the sharpening of the boundary signals, as shown in figure 3.
Figure 3: The spatial response profiles for the strip and the square figures 100 msec after stimulus onset show response peaks right at the center (0°) as well as at the boundaries (±2°) of the 4° wide strip and square. The abscissa is the distance in visual angle from the RF to the center of the strip or the square.

For a "negative" stimulus in which the texture within the figure is orthogonal to a cell’s preferred orientation, the cell tends to respond initially only to the background texture but not to the figure. However, after an initial silence, a central peak would sometimes develop as the boundary signal sharpens, as shown in Figure 4.

Figure 4: A central peak emerged gradually together with the sharpening of the boundary signals in a region where there was no initial response to the local stimulus.

While the boundary signal sharpens, the responses within the figure’s interior also become more smooth over time and finally reach a uniform level, which is independent of the size of the figure, as shown in Figure 5.
Figure 5: A multiunit response to boundaries generated by squares of five different sizes (3", 4", 5", 6", 8" in width) exhibit a similar gradual sharpening effect. The centers of all the squares are located at 0". In this example, at 80 msec, the responses of the unit inside the figure are inversely proportional to the figure size (see also [6]). By 140 msec, the responses within the interior of the figures have been “smoothed” to become more uniform.

A figure-ground enhancement ratio can be computed by dividing the response of the cell when it is located within the figure by the response of the cell to identical stimulus when it is located outside the figure for each position. The spatial profiles of the enhancement ratios often exhibit the “batman” profile shown in Figure 6.

Figure 6: The “figure-ground” enhancement profile, computed for the period 100-200 msec after stimulus onset, exhibits a central enhancement peak (0") and enhancement peaks right inside the boundaries (±2"). The medial peak might signify the medial axis whereas the peak inside the boundary might play a role in signaling border-ownership.

Summary and Conclusions

Our results show that V1 neurons can detect figure boundaries defined by texture contrast (see also [7]). The initial responses of the cells are specific to the local stimuli. Between 60-120 msec, the responses are characterized by the gradual emergence and sharpening of the boundary signals. This is accompanied by a simultaneous suppression and smoothing of the signals within the surface. The 100 msec required for sharpening the boundary signals is consistent with the psychophysical time required for perceiving texture segmentation [8].

The enhancement profile within the square stimulus is consistent with Lamme’s [5] findings that V1 neurons produce stronger responses when they are inside a figure than when they are outside. Our results revealed that the enhancement profile is characterized by important spatial structures which are similar to those in the perceptual sensitivity profile observed by Kovacs and Julesz [9] in their psychophysical experiments. The central peaks in these profiles have been
suggested to play a role in representing the medial axis (i.e. axis of symmetry) of objects.

These findings provide evidence that V1 neurons not only participate in the dynamic processes of image segmentation, but could also involve in signaling the axis of symmetry of objects. The spatio-temporal response and enhancement profiles suggest that individual V1 neurons might be able to extract, represent and signal multiple types of information, including the boundary, surface qualities and the axis of symmetry of objects in a visual scene.

Acknowledgement: This research was conducted in the laboratory of Peter Schiller in MIT and was funded by EY00676 to Peter Schiller and a NSF grant DMS-93-21266 to David Mumford at Harvard University. The author is grateful to Victor Lamme, Karl Zipser, Edward Tehovnik, Marc Sommer, J-Han Chou, Andreas Tolias, James Mazer, Tim Moore, Warren Slocum, Peter Schiller and David Mumford for insightful discussion and technical guidance.

References