Investigating Shape Representation in Area V4 with HMAX

Minjoon Kouh and Maximilian Riesenhuber

The Problem: The ventral pathway of the primate cortex is thought to mediate object recognition. Within the early visual areas along the pathway, neurons tend to respond well to oriented bars or edges. Neurons in the intermediate visual areas are no longer tuned to the oriented bars only, but to other forms and shapes of intermediate complexity. Finally in the high visual areas, such as the inferior temporal cortex (IT), neurons are responsive to complex shapes like the image of a face or a hand [2] [3] [4]. However, the neural mechanism by which such progression of selectivity arises is an open question.

Motivation: Understanding how the neuronal population represents shape information is one of the main objectives of visual neuroscience. A computational theory can provide testable hypotheses to analyze existing data with and to motivate further experiments.

Previous Work: We have recently presented a computational model of object recognition in cortex [1]. The model is composed of a hierarchy of feed-forward layers of neuron-like units, performing either one of two operations: A weighted linear sum (template match) to increase feature complexity, and a nonlinear pooling operation based on a maximum operation to increase response invariance to translation and scaling.

In preliminary work, we have compared the shape tuning of units in intermediate layers in HMAX to experimental data from several experiments. In particular, it has been shown that, with appropriately chosen model parameters, the orientation tuning and the bandwidths of the model units lie within a physiologically plausible range as reported in [2] and [3]. (See Table 1.) The model units are also tuned to multiple classes of gratings (Cartesian, polar, and hyperbolic) and generally biased toward non-Cartesian stimuli as reported in [4]. (See Table 2.)

<table>
<thead>
<tr>
<th>Median Orientation Bandwidths</th>
<th>Physiology</th>
<th>HMAX</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>V1</td>
<td>V4</td>
</tr>
<tr>
<td>All neurons</td>
<td>42°</td>
<td>75°</td>
</tr>
<tr>
<td>Less than 90°</td>
<td>37°</td>
<td>52°</td>
</tr>
</tbody>
</table>

Table 1: Median orientation bandwidths from physiology (V1, V4) and simulation results (S1, S2). The physiology data were taken from [2] for V1 and [5] for V4. The model data are from [8] and [9].

<table>
<thead>
<tr>
<th>Physiology</th>
<th>HMAX</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polar</td>
<td>10%</td>
</tr>
<tr>
<td>Hyperbolic</td>
<td>8%</td>
</tr>
<tr>
<td>Cartesian</td>
<td>2%</td>
</tr>
</tbody>
</table>

Table 2: Percentage of cells that gave more than twice the peak responses to one stimulus class than to another. The physiology data were taken from [4]. The model data are from [8].

Approach: The ultimate goal is to understand how shape representation progresses along the ventral pathway. This unknown will be probed with a computational model, by using the same sets of stimuli and the experimental procedures as in the physiological studies and by comparing the experimental data with the simulation.

The first hypothesis we will investigate is whether the experimental data on V4 neuron shape tuning can be explained as a result of combining complex cell-like inputs. For instance, Pasupathy and Connor have recently shown that neurons in area V4 appear to be tuned to contour features, such as curvature and convexity [5]. They have further postulated that V4 neurons might represent not just local shape but also implement some form of object-centered coding [6] [7]. Using the HMAX model and the experimental stimuli and paradigm, our research efforts will focus on comparing the tuning of model units to those of experimental neurons in order to understand the computational processes involved.

Difficulty: While the shape tuning properties of neurons in the early visual pathway, such as area V1, have been relatively well characterized, there has been very little theoretical work for the higher visual areas, due to their nonlinear behavior. The wide variety of stimuli employed by different experimental groups — with each group reporting some tuning of V4 neurons to the stimuli used in their study — is testament to this problem and points to the need for a computational model of shape representation in cortex to integrate the data and generate testable a priori hypotheses.
**Impact:** This study investigates how visual information (shape and form) is represented in visual cortex. Ultimately, such understanding will provide better insight into object recognition and classification processes. Moreover, building more complex tuning from afferents with simpler tuning is a crucial computational operation, and insight into the computational mechanisms underlying it will likely be of relevance also for understanding processing in other brain areas, including other sensory modalities.

**Future Work:** Once we have achieved a better theoretical understanding of the existing data regarding shape representation in area V4 and have integrated the results into our model, one goal is to use these insights to make predictions for further experiments, and to extend our analysis to other brain areas.

**Research Support:** This report describes research done at the Center for Biological & Computational Learning, which is in the Department of Brain & Cognitive Sciences at MIT and which is affiliated with the McGovern Institute of Brain Research and with the Artificial Intelligence Laboratory.

This research was sponsored by grants from: Office of Naval Research (DARPA) Contract No. N00014-00-1-0907, Office of Naval Research (DARPA) Contract No. N00014-02-1-0915, National Science Foundation (ITR/IM) Contract No. IIS-0085836, National Science Foundation (ITR/SYS) Contract No. IIS-0112991, and National Science Foundation-NIH (CRCNS) Contract No. IIS-0085836.


**References:**


