Rethinking how the Nervous system registers and identifies shapes

Ernest Greene

University of Southern California, USA

Most current theories of shape recognition propose that lines and edges (contours) are registered as elemental building blocks, with identification of a shape being accomplished by an enumeration of those elements. The critical role of contours has been assumed by artists and philosophers for centuries, and by psychologists since the inception of our discipline. Many of the early concepts were not clearly delineated, but the advent of computers provided a better understanding of how to specify the task as mechanistic steps. Both Selfridge (1959) and Marr (1982) advanced conceptual models for how the lines and edges of shapes could be registered and then combined to allow for identification. Further, the Nobel Prize winning research of Hubel & Wiesel (1959; 1962) appeared to provide neural substrates for registering contour attributes, and a plausible theory for how those attributes could be combined by the nervous system. These investigators found that individual neurons in primary visual cortex (V1) were selectively activated by elongated bars, with the degree of activation being determined by the orientation and location of a given stimulus. Their model for how the neurons manifested this selectivity was based on precise anatomical mapping of connections from retina, through lateral geniculate nucleus, to cortex, wherein an aligned set of retinal ganglion cells provided the stimulation to a given "orientation selective" neuron in V1. The various contours comprising a given shape would be expected to drive activity in a specific subset of V1 neurons. It seemed plausible that the output from neurons in that subset could converge on a higher order neuron, such that the receiving neuron would be activated only by the contours of that particular shape.

The theory that shape-recognition mechanisms are based on anatomical convergence of axons has often been challenged as implausible on the basis of the sheer number of shapes that humans are able to recognize. That number expands exponentially as one must account for variations of size and retinal location. The capsule phrase for this criticism is known as the "combinatorial explosion," which highlights the point that the model requires a specific combination of anatomical connections to provide for recognition of each identifiable shape, at each angel of view, variation of size, and variation of location within a briefly glimpsed scene.

Further, there is reason to doubt the initial proposition that contour segments are registered as elemental shape components, and more specifically, that the orientation, curvature, and linear extent of the contours are critical features for defining a given shape. A shape can be rendered as sequence of discrete dots that mark the boundaries of the shape, similar to a silhouette. One can then reduce the density of the dots being displayed, for example by showing only every fourth dot, or every sixth dot, and then asking observers to name the shape. Most are able to name diverse shapes that display only every 10th dot, and many shapes can be identified at an even lower density (Greene, 2007). It is not plausible that orientation-selective neurons are being activated by these sparse dot patterns. Shapes were identified even when the spacing between two adjacent dots was larger than the receptive field of orientation-selective neurons. Further, many dots on opposite sides of the shape were closer together than were the dots that followed the contour. Therefore one could not use proximity to "reconstruct" the outer boundary of the shape. There is, at present, no known neurophysiological mechanism for how dot patterns can elicit identification of diverse shapes.

Other work from this lab provides an even greater challenge to the concept that contours are elemental shape features (Greene & Hautus, 2017). One can briefly display an unknown (random) shape that consists of a continuous string of dots that constitute an outer boundary. This one-time exposure is followed by a low density version of the same shape, or a low density version of a different shape, requiring that the observer say whether or not the second display was derived from the first. The challenge is made more demanding by displaying the low-density option at a different location than where the initial target was shown. Under these conditions, a 12-15% dot density can provide for very high levels of correct responding, and a 3% density is still above chance.

Studies such as these provide evidence that shapes can be identified from sparse dot patterns, and intact contours are not elemental features that define a given shape. The extended lines and edges that serve as boundaries and internal features of objects appear to be highly over-determined, with each segment providing a congeries of marker locations that contribute to shape identification. It appears that the mechanisms required for shape recognition are the same as, or are very similar to, those needed for pattern recognition.

One might note that computer-based theories for pattern recognition require an address for each marked location. Unfortunately, there is at present no neurophysiological principle for specifying the address of a stimulated location – another weighty problem that may be discussed at a later time.

References

Greene, E. (2007) Recognition of objects displayed with incomplete sets of discrete boundary dots. Perceptual and Motor Skills, 104, 1043-1059.

Greene, E. and Hautus, M.J. (2017) Demonstrating invariant encoding of shapes using a matching judgment protocol. AIMS Neuroscience, 4, 120-147.

Hubel, D.H. and Wiesel, T.N. (1959) Receptive fields of single neurons in the cat's striate cortex. Journal of Physiology, 148, 574-591.

Hubel, D.H., Wiesel, T.N. (1962) Receptive fields, binocular interactions and functional architecture in the cat's visual cortex. Journal of Physiology, 160, 106-154.

Marr D. (1982) Vision: a computational investigation into the human representation and processing of information. New York: Freemen, pp. 51-79.

Selfridge, O.G. (1959) Pandemoniium: a paradigm for learning. In. D.V. Blake, A.M. Uttley (Eds.) Symposium on the Mechanization of Thought Processes, pp. 103-158. London: H.M. Stationary Office.