

New mathematical model of brain information processing predicts some of vision peculiarities

28 January 2011, by Larry Hardesty



Graphic: Christine Daniloff

The human retina -- the part of the eye that converts incoming light into electrochemical signals -- has about 100 million light-sensitive cells. So retinal images contain a huge amount of data. High-level visual-processing tasks -- like object recognition, gauging size and distance, or calculating the trajectory of a moving object -- couldn't possibly preserve all that data: The brain just doesn't have enough neurons. So vision scientists have long assumed that the brain must somehow summarize the content of retinal images, reducing their informational load before passing them on to higher-order processes.

At the Society of Photo-Optical Instrumentation Engineers' Human Vision and Electronic Imaging conference on Jan. 27, Ruth Rosenholtz, a principal research scientist in the Department of Brain and Cognitive Sciences, presented a new [mathematical model](#) of how the brain does that summarizing. The model accurately predicts the

visual system's failure on certain types of image-processing tasks, a good indication that it captures some aspect of human cognition.

Most models of human [object recognition](#) assume that the first thing the brain does with a retinal image is identify edges -- boundaries between regions with different light-reflective properties -- and sort them according to alignment: horizontal, vertical and diagonal. Then, the story goes, the brain starts assembling these features into primitive shapes, registering, for instance, that in some part of the visual field, a horizontal feature appears above a vertical feature, or two diagonals cross each other. From these primitive shapes, it builds up more complex shapes - four L's with different orientations, for instance, would make a square - and so on, until it's constructed shapes that it can identify as features of known objects.

While this might be a good model of what happens at the center of the visual field, Rosenholtz argues, it's probably less applicable to the periphery, where human object discrimination is notoriously weak. In a series of papers in the last few years, Rosenholtz has proposed that cognitive scientists instead think of the brain as collecting [statistics](#) on the features in different patches of the visual field.

Patchy impressions

On Rosenholtz's model, the patches described by the statistics get larger the farther they are from the center. This corresponds with a loss of information, in the same sense that, say, the average income for a city is less informative than the average income for every household in the city. At the center of the visual field, the patches might be so small that the statistics amount to the same thing as descriptions of individual features: A 100-percent concentration of horizontal features

could indicate a single horizontal feature. So Rosenholtz's model would converge with the standard model.

But at the edges of the visual field, the models come apart. A large patch whose statistics are, say, 50 percent horizontal features and 50 percent vertical could contain an array of a dozen plus signs, or an assortment of vertical and horizontal lines, or a grid of boxes.

In fact, Rosenholtz's model includes statistics on much more than just orientation of features: There are also measures of things like feature size, brightness and color, and averages of other features - about 1,000 numbers in all. But in computer simulations, storing even 1,000 statistics for every patch of the visual field requires only one-90th as many virtual neurons as storing visual features themselves, suggesting that statistical summary could be the type of space-saving technique the brain would want to exploit.

Rosenholtz's model grew out of her investigation of a phenomenon called visual crowding. If you were to concentrate your gaze on a point at the center of a mostly blank sheet of paper, you might be able to identify a solitary A at the left edge of the page. But you would fail to identify an identical A at the right edge, the same distance from the center, if instead of standing on its own it were in the center of the word "BOARD."

A + B O A R D

Rosenholtz's approach explains this disparity: The statistics of the lone A are specific enough to A's that the brain can infer the letter's shape; but the statistics of the corresponding patch on the other side of the visual field also factor in the features of the B, O, R and D, resulting in aggregate values that don't identify any of the letters clearly.

Road test

Rosenholtz's group has also conducted a series of experiments with human subjects designed to test the validity of the model. Subjects might, for instance, be asked to search for a target object - like the letter O - amid a sea of "distractors" - say, a jumble of other letters. A patch of the [visual field](#) that contains 11 Q's and one O would have very similar statistics to one that contains a dozen Q's. But it would have much different statistics than a patch that contained a dozen plus signs. In experiments, the degree of difference between the statistics of different patches is an extremely good predictor of how quickly subjects can find a target object: It's much easier to find an O among plus signs than it is to find it amid Q's.

Rosenholtz, who has a joint appointment to the Computer Science and Artificial Intelligence Laboratory, is also interested in the implications of her work for data visualization, an active research area in its own right. For instance, designing subway maps with an eye to maximizing the differences between the summary statistics of different regions could make them easier for rushing commuters to take in at a glance.

In vision science, "there's long been this notion that somehow what the periphery is for is texture," says Denis Pelli, a professor of psychology and neural science at New York University. Rosenholtz's work, he says, "is turning it into real calculations rather than just a side comment." Pelli points out that the [brain](#) probably doesn't track exactly the 1,000-odd statistics that Rosenholtz has used, and indeed, Rosenholtz says that she simply adopted a group of statistics commonly used to describe visual data in computer vision research. But Pelli also adds that visual experiments like the ones that Rosenholtz is performing are the right way to narrow down the list to "the ones that really matter."

Provided by Massachusetts Institute of Technology

APA citation: New mathematical model of brain information processing predicts some of vision peculiarities (2011, January 28) retrieved 20 September 2019 from <https://medicalxpress.com/news/2011-01-mathematical-brain-vision-peculiarities.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.