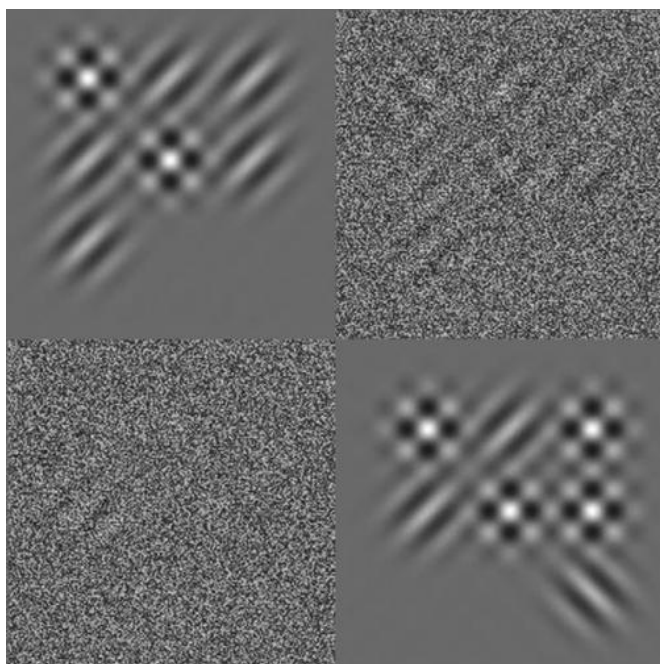# Step by step: Feature detection and combination in perceptual learning and object identification

11 January 2013, by Stuart Mason Dambrot

Stimuli. (Upper Left) A Gabor letter. When unconstrained, the human participant is presented with a Gabor letter faintly in noise (Upper Right). As the detector, the participant is presented with a single feature faintly in noise (Lower Left) and, as the combiner, with an imperfect set of detected features (Lower Right). In this last case, the high-contrast Gabors are easily seen, but are a less-than-faithful copy of the original letter's features, which makes it hard to guess what the original letter was. Copyright © PNAS, doi:10.1073/pnas.1218438110

(Medical Xpress)—The ease and immediacy with which we recognize familiar objects escapes our notice. However, a novel, ambiguous, or highly complex object requires practice to achieve such perceptual facility. Past perceptual learning research found a wide range of rates at which these object recognition skills are acquired.

Recently, however, scientists at Harvard University and New York University have devised a way to distinguish feature detection and feature combination, and moreover have determined the rate at which these two steps improve during perceptual learning. The researchers found that while detection is inefficient and learned slowly, combination is learned at a rate four to seven times greater. In addition, they show how this clarifies the diverse results obtained in previous perceptual learning studies.

NYU Professor Denis G. Pelli and Harvard PhD student Jordan W. Suchow faced a number of challenges in conducting their research. These include devising a method to separate detection and combination, and reveal how each improves as the observer learns. "It's widely supposed that object recognition proceeds in two steps," Suchow tells Medical Xpress. "First the viewer detects basic features such as line segments and patches of dark and light, and then the viewer combines the features to form the object. Viewers do both steps every time they identify an object, and it's impossible for them to voluntarily turn one step off, using only the other."

Measuring the steps separately, notes Suchow, required a new trick. "It was a great thrill when we found that the efficiency of unconstrained human performance was predicted by the product of the separately measured efficiencies of the two steps," Pelli adds.

To solve this, Suchow explains, the scientists designed a pair of *bionic* crutches, each a computer program that does one of the steps optimally (that is, as accurately as possible). "By having the human do the detecting and the computer do the combining, or vice versa, we can separately measure each step. Because the computer

[program](link) is optimal, if the human and computer together perform less than optimally, we know that it's the human's fault."

Pelli also relates their feature identification findings to the neurobiology of feature extraction of visual primitives, as first identified by David Hubel and Torsten Wiesel. In their groundbreaking 1959 paper[1] Hubel and Wiesel showed that what they termed *simple cells* in the primary visual cortex act as feature detectors – that is, they cross-correlate the image with a known signal within a region of space referred to as a *receptive field profile*. "It has long been known that the brain must, somehow, combine the activity of several feature detectors to do typical object identification tasks," Pelli explains. "However, the rules of combining are still mysterious. Our results confirm the conjecture that the identification process can be modeled by two steps – detecting and combining. Our method traces the very different learning trajectories of the two neural parts of the architecture."

Regarding face specificity in the brain's fusiform face area, Pelli adds, "Our results apply to object recognition in general and letter identification in particular. It would be interesting to do new experiments targeting face recognition in particular."

The bionic crutches – ideal detector and ideal combiner – were key components of the study. "This is the hybrid of two distinct traditions," says Pelli. "Signal detection theory specifies the optimal algorithm, while feature detection posits a seemingly dumb rule for how people identify, namely in two steps. The innovation was to provide, as an aid, optimal solutions for each step to the human." The bionic crutches made it possible for Pelli and Suchow to assess the efficiency of the human subject's performance in each step.
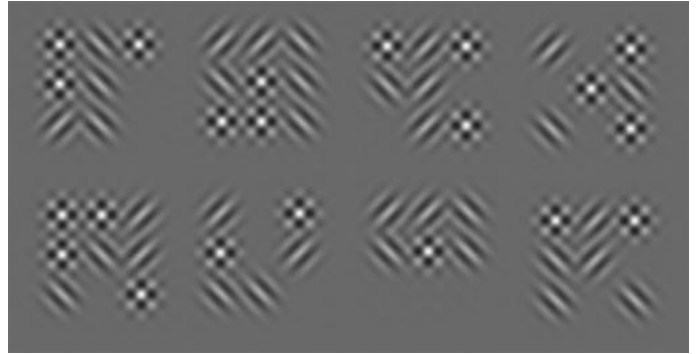
Eight Gabor letters. The letters of the IndyEighteen alphabet are composed of Gabors. Each of the 18 possible Gabors is oriented ±45° from vertical and is at one of nine locations in a 3 × 3 grid. When a right-tilted and a left-tilted Gabor coincide, they form a plaid, but vision still responds to them independently. We suppose that the Gabors are detected independently, so that each Gabor is a feature. With two orientations and nine locations, there are 18 possible Gabors, i.e., features. The eight letters displayed here are a randomly selected subset of the 218 letters in the whole alphabet. Note that within this subset, some features are common to many letters (e.g., six of the eight letters contain a right-tilted Gabor at the top right corner), whereas some features are common to just a few (e.g., two of the eight letters contain a right-tilted Gabor at the bottom left position). Copyright © PNAS, doi:10.1073/pnas.1218438110

Also significant is that the study findings (that detection is inefficient and learned slowly, and combining at an increasingly faster rate) explain much of the diversity of rates reported in [perceptual learning](link) studies – especially regarding effects of complexity and familiarity. "With practice," notes Suchow, "people learn to see better, recognizing objects that were once too small, faint, distorted, or unfamiliar. This takes practice, and the amount of practice that's needed depends on the task."

With simple tasks – like detecting a faint flicker on a computer screen, he illustrates – even many weeks of practice yield little improvement. "We think that's because those tasks rely primarily on the detection mechanism, which learns slowly." With other tasks – such as identifying a foreign letter – there's a big improvement and it happens very quickly. "We think that's because those tasks are more dependent on learning to combine the features, and the

combining mechanisms learn quickly."

In terms of next steps, Suchow notes that their experiments relied on a specially-constructed alphabet, comprised of so-called *Gabor letters*, which is similar in appearance to Braille but has never before been seen or used. "It would be great to extend the method beyond this alphabet to other sets of objects, such as faces," he adds, agreeing with Pelli.

Suchow also sees possible applications of their work to other areas of research. "Seeing is usually effortless, but, again, there are important cases where it's hard – for example, when the objects are small, faint, distorted, or unfamiliar. In these cases, understanding which stages of vision limit our ability to perceive and learn may guide the creation of new technologies and designs that help us to see."

  More information: Learning to detect and combine the features of an object, *PNAS* published online before print December 24, 2012, doi:10.1073/pnas.1218438110

Related:
[1] [Receptive fields of single neurones in the cat's striate cortex](), *Journal of Physiology* 1959 October; 148(3): 574–591